



L'Intelligenza Artificiale Generativa nella Giustizia Amministrativa: scenari, rischi e opportunità

Germana Lo Sapio
3 febbraio 2025

There's a way to do it better. Find it.
Thomas Edison

Sommario*

1. Lo scenario
2. Dove LLM e Giustizia si incontrano: linguaggio e digitalizzazione
3. “*Correlation is not causation*”: un mito da sfatare o un limite destinato a durare?
4. La Giustizia nell’AI Act e le vie di fuga
5. I Rischi dei LLM applicati alla Giustizia e le possibili misure di mitigazione
 - 5.1. Il rischio di *automation bias*
 - 5.2. Il rischio di allucinazioni
 - 5.3. Il rischio di omologazione alla cultura giuridica di *Common Law*
6. Orizzonti applicativi dei LLM nella Giustizia Amministrativa
 - 6.1. La ricerca “intelligente”
 - 6.2. Analisi dei documenti e degli atti del fascicolo digitale
 - 6.3. Chiarezza e sinteticità degli atti con il supporto di IA
 - 6.4. La *vis* espansiva di IA nelle ADR
7. Conclusioni: il rischio inedito della *Shadow AI*

1. LO SCENARIO

Abstract¹: I Large Language Model (LLM), come ChatGPT di OpenAI lanciato nel 2022, hanno fatto conoscere al mondo l’Intelligenza Artificiale Generativa, offrendo prestazioni linguistiche paragonabili a quelle umane, con velocità e adattabilità senza precedenti. Progettati per generare testi in qualsiasi lingua e stile, i LLM hanno catalizzato l’innovazione tecnologica, sollevando questioni su etica, privacy e rischi per gli ordinamenti democratici. La loro capacità di operare con un “certo grado di autonomia” li distingue dalle innovazioni del passato e li pone al centro di una trasformazione che interessa anche il mondo giuridico, inducendo ad interrogarsi sul se e come il mondo del diritto potrà essere trasformato dall’IA stessa.

I “**Large Language Model**” (LLM)² incarnano l’avanguardia dell’**Intelligenza Artificiale** oggi, che è soprattutto l’**IA Generativa**³. Ma nel mondo dell’IA, spinto dalla velocità

**Il presente articolo è la rielaborazione aggiornata della relazione per il Congresso nazionale dei magistrati amministrativi italiani, I 150 anni di funzionamento dei Tribunali amministrativi regionali, che si è svolto a Roma il 18 e 19 ottobre 2024.*

¹L’abstract dei singoli paragrafi è stato generato da ChatGPT (versione GPT 4.0). E’ stata inserita nel sistema la copia del paragrafo ed è stata formulata questa domanda (prompt): “Fai un abstract di questo paragrafo che sia massimo di 10 righe e che contenga anche qualche informazione specifica riportata nel paragrafo stesso”. Contrariamente a ciò che correttamente dovrebbe avvenire nell’interazione con un sistema basato su un LLM, ho deciso di non modificare il testo proposto da ChatGPT e di formulare un prompt molto generico. Ogni risposta è stata generata in circa 2 secondi di tempo. Le locuzioni scritte in carattere **grassetto** sono oggetto del **Glossario**, in appendice.

² Nel testo verrà usato l’acronimo LLM, ma i più avanzati modelli di grandi dimensioni si sono già trasformati in “multimodali”, cosicché la corretta locuzione sarebbe **LMM**, **Large Multimodal Model**: essi hanno infatti l’abilità di poiché possono elaborare e generare contestualmente video, immagini, musica, linguaggio di programmazione. I benefici attesi dall’IA nel prossimo futuro ruotano proprio intorno al mondo delle immagini, soprattutto nel contesto sanitario: basti pensare allo screening delle immagini per le diagnosi mediche, <https://www.sanitainnovazione.digitalizzazione.it/2024/07/06/caso-eclatante-e-innovativo-google-health-e-la-diagnosi-del-cancro-al-seno/>. Quanto alle abilità eccezionali dei LLM nella elaborazione anche del linguaggio di programmazione: cfr., P. L. Pisa, Sundar Pichai, Ceo di Google: *Ora l’IA scrive più di un quarto del nostro nuovo codice*, in La Repubblica, 30 ottobre 2024.

³ J. Kaplan, *Generativa AI. Conoscere, capire e usare l’Intelligenza Artificiale generativa*, ed. italiana Luiss University Press, 2024, p.15. Per i filosofi e gli scienziati, l’IA Generativa (*Gen.IA*) è una tecnologia che divide la Storia dell’umanità in due, con un *prima* e un *dopo* il suo avvento (cd. Periodizzazione), tanto è pervasiva in ogni settore della cultura, dell’economia, della società, inducendo a ripensare la stessa dimensione del rapporto tra uomo e tecnica; cfr. R. Manzotti, S. Rossi, *Io&la mente. Cervello e GPT*, Soveria Mantelli Rubbettino, 2023, n. p. 137.

travolgente dell'innovazione tecnologica, ciò che appare rivoluzionario al momento presente è il preludio di ciò che domani prenderà il suo posto⁴.

I LLM sono modelli⁵ di IA progettati e sviluppati per comprendere e generare testi in linguaggio *naturale* e, in molti documenti di *policy* governative, sono indicati come i “*modelli di frontiera dell'IA*”⁶. Nel campo delle scienze computazionali, per una macchina comprendere e rielaborare il linguaggio naturale è stato, fin dagli anni '50, il “*Santo Graal*” della ricerca. La branca che se ne occupa va sotto il nome di *Natural Language Processing* (NLP). Per decenni, però i risultati non sono stati brillanti, come può confermare chiunque, fino a pochi anni fa, si sia cimentato con le applicazioni di IA di traduzione automatica come “*Google Translator*” o con le prime versioni di *Siri* o *Alexa* per il riconoscimento vocale. I progressi più recenti, che ora sono sotto gli occhi di tutti, sono stati determinati da invenzioni geniali nell'architettura algoritmica dei modelli, dall'enorme mole di dati digitali rinvenibili grazie al *Social Web* (cd. Web 2.0) sui quali sono addestrati i LLM, e, dalla potenza computazionale delle più moderne infrastrutture di calcolo e di chip avanzati, sul cui reperimento è in atto una vera e propria “*corsa all'oro*”, in Cina già ribattezzata come la “*guerra dei 100 modelli*”⁷.

Tuttavia, nella percezione comune, la svolta è sembrata improvvisa e ha una data di inizio: il 30 novembre 2022, giorno del lancio, tramite interfaccia web, di **ChatGPT**, rilasciato gratuitamente da OpenAI.

⁴ J. Kaplan, *Generativa AI. Conoscere, capire e usare l'Intelligenza Artificiale generativa*, cit., pag. 33 “*La tecnologia informatica avanza a passi da gigante, un fenomeno descritto per la prima volta già nel 1965 da Gordon Moore, cofondatore di Intel, tramite quella che oggi chiamiamo “Legge di Moore” (...) Moore spiegò che, a partire dall'invenzione del circuito integrato, la densità dei transistor sui chip e altri indicatori di potenza dei computer come la velocità e la memoria erano raddoppiati ogni anno e mezzo. In modo davvero stupefacente, questa tendenza è andata avanti più o meno per mezzo secolo. Ci troviamo di fronte ad un esempio di crescita esponenziale, con un concetto per molti difficile da comprendere (...) La questione è cruciale: a un certo punto le differenze quantitative divengono qualitative*”. Per alcuni studiosi, le eccezionali abilità linguistiche dei LLM sono proprio dovute anche all'enorme scala di potenza di calcolo, all'uso di piattaforme hardware dedicate come le **GPU (Graphic Processing Units)** e alla mole di dati utilizzati per l'addestramento, resa possibile dal dilagare di **Social Web**.

⁵ Il termine “*modello*” è utilizzato anche dal legislatore europeo nell'AI Act (Regolamento 2024/1689 UE) per distinguerlo dai “*sistemi di IA*”. La definizione di “*modello con finalità generale*”, dentro cui rientrano anche i LLM, è contenuta nell'art. 3 n. 63: “*un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato da una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione, prima di essere immessi sul mercato*”. La definizione di sistema di IA è invece contenuta, come noto, nell'art. 3 n. 1 dell'AI Act. “*un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali*”. I “*modelli di IA*” e “*sistemi di IA*” sono destinati ad integrarsi nella pratica, considerando che i primi fungono da base per i secondi, come attestato dallo stesso più noto ChatGPT in cui il modello è GPT e il sistema è il chatbot. Per un primo esame dell'approccio regolatorio sul tema e delle distinzioni tra le varie categorie utilizzate nel regolamento, sia consentito il rinvio a G. Lo Sapio, *L'Artificial Intelligence Act e la prova di resistenza per la legalità algoritmica*, Federalismi.it, 10 luglio 2024.

⁶ Era incentrato sull'IA Generativa e i LLM venivano menzionati come “*frontier model*” nel *Memorandum sulla Sicurezza Nazionale* adottato dal Governo degli Stati Uniti, il 24 ottobre 2024, adottato in attuazione dell'*Executive Order* di Biden del 30 ottobre 2023 sullo sviluppo e l'uso sicuro, affidabile e responsabile dell'intelligenza artificiale. Come noto, il 23 gennaio 2025, l'*Executive Order* di Biden è stato revocato da un *Executive Order* di Trump, al dichiarato scopo di eliminare inutili barriere per il rafforzamento della leadership degli Stati Uniti nell'IA. La sezione 1, recita “*The United States has long been at the forefront of artificial intelligence (AI) innovation, driven by the strength of our free markets, world-class research institutions, and entrepreneurial spirit. To maintain this leadership, we must develop AI systems that are free from ideological bias or engineered social agendas. With the right Government policies, we can solidify our position as the global leader in AI and secure a brighter future for all Americans. This order revokes certain existing AI policies and directives that act as barriers to American AI innovation, clearing a path for the United States to act decisively to retain global leadership in artificial intelligence*”.

⁷ <https://business.cornell.edu/hub/2024/12/13/chinas-domestic-ai-competition-heats-up/>

Dal quel giorno in poi, i riflettori, anche dei legislatori di tutto il mondo, si sono accesi e la scena è cambiata nel giro di pochi mesi⁸. Le prestazioni linguistiche di questo chatbot, poi seguito da modelli sempre più avanzati delle altre grandi “*Compagnie del digitale*”⁹, hanno rivelato qualità mai viste prima: coerenza linguistica, correttezza grammaticale, fluidità discorsiva, in qualunque lingua, con risultati sostanzialmente indistinguibili da quelli dei testi scritti da esseri umani. Un successo planetario, ottenuto all’inizio anche grazie al passaparola, fondato su due elementi chiave: l’enorme facilità d’uso e la amplissima sfera di applicazione, che ha consentito agli utilizzatori di sbizzarrirsi con anche la propria creatività.

Negli ultimi due anni, il dibattito sui rischi e sulle opportunità dell’IA, anche nel dominio giuridico, è stato serrato: sono stati pubblicati centinaia di articoli su riviste scientifiche, organizzati convegni, *online* e in presenza, redatti report, promulgate Carte di principi¹⁰ - da ultimo anche di ordini professionali forensi¹¹- pubblicati libri che spaziano dal taglio giuridico, a quello tecnico, a quello filosofico, o storico-politico.

Una pletera di materiale che meriterebbe un LLM dedicato per potere essere esaminato e sistematizzato. In gioco ci sono quesiti esistenziali sul destino dell’Umanità nel rapporto con queste nuove forme di “*intelligenza aliena*” (“*Non-Human-Intelligence*”), ma anche domande pragmatiche, allo stato prive di risposte certe: quale sarà il reale impatto dei LLM nel mercato del lavoro¹², sul ruolo e sulle competenze dei giuristi e nella giustizia in generale.

⁸ “Prima di ChatGPT ciò che i profani sapevano e pensavano dell’IA veniva da film di fantascienza, articoli di giornale o servizi televisivi, ma non c’era la sensazione che l’IA facesse già parte della nostra vita. Con L’avvento di ChatGPT, invece, per la prima volta nella storia dell’IA tutti pensano che sia improvvisamente arrivata, che abbia fatto il suo ingresso trionfale o catastrofico, a seconda dei punti di vista, nella nostra quotidianità” F. Rossi, *Intelligenza Artificiale. Come funziona e dove ci porta la tecnologia che sta trasformando il mondo*, Laterza, 2024, p. 18. Irrompendo sulla scena a metà del percorso, i LLM e l’IA Generativa hanno catalizzato l’attenzione anche dei co-legislatori dell’Unione Europea, intanto alle prese con i triloghi relativi dell’IA Act, oggi Regolamento 2024/1689 UE. Un primo timido tentativo di inglobare nell’AI Act anche i cd. *modelli con finalità generali*, si rinviene solo nel testo di compromesso del 29 novembre 2021 che prevedeva l’art. 52a e il connesso *Consideration 70*. Ma dopo il lancio di ChatGPT a novembre 2022, il percorso legislativo su come e se regolare i modelli con finalità generali, oggi disciplinati dal Capo V del regolamento 2024/1689, è stato alquanto tumultuoso anche per le posizioni contrarie di Italia, Francia e Germania, fino a confluire nel “*record assoluto di 56 ore di continuative di negoziato*” tra Parlamento, Commissione europea e Consiglio durato dal pomeriggio del 6 dicembre 2023 fino alla notte del 9 dicembre 2023. Cfr. R. Viola, L. De Biase, *La legge dell’Intelligenza Artificiale. L’AI Act per trovare un senso nel futuro della tecnologia*, Il Sole24ore, 2024, pag. 134 e ss

⁹ Luciano Violante ha coniato questa efficace locuzione che riecheggia quella della “*Compagnia Inglese delle Indie Orientali*”, per descrivere le Big Tech dell’oligopolio dell’IA (Microsoft, Google, Amazon, Meta, OpenAI) che governano il mercato del cloud infrastrutturale e le caselle di posta elettronica, solo che, a differenza delle antiche Compagnie delle Indie, non colonizzano territori inesplorati, ma creano esse stesse il territorio che occupano, ossia lo spazio digitale. Cfr. O. Pollicino, P. Dunn, *Intelligenza artificiale e democrazia. Opportunità e rischi di disinformazione e discriminazione con Prefazione* di L. Violante, Bocconi University Press, 2024.

¹⁰In merito, pare opportuno segnalare che il Gruppo di lavoro sulla giustizia cibernetica e sull’intelligenza artificiale (CEPEJ-GT-CYBERJUST), istituito nell’ambito del Cepej (Commissione Europea per l’Efficienza della Giustizia) del Consiglio d’Europa, a febbraio 2024, ha adottato la nota “*Use of Generative Artificial Intelligence (AI) by judicial professionals in a work related context*”, proprio con specifico riguardo all’AI Generativa, sottolineando le aree dell’attività giurisdizionale in cui tali presenterebbero “*buoni risultati all’interno di un contesto chiaramente definito*”: traduzione di testi; generazione di testi, immagini o suoni coerenti; riassunto automatico di testi, analisi semantica e rilevamento delle opinioni, *text mining* e accesso ai contenuti. Proprio il CEPEJ, nel 2018, per la prima volta ha accostato i termini “giustizia” e Intelligenza artificiale, nel medesimo anno in cui le Istituzioni dell’Unione Europea istituivano il Gruppo di Alti Esperti in materia: cfr. *European Ethical Charter on the Use of Artificial Intelligence in Judicial System and their Environment*, 4 dicembre 2018.

¹¹ <https://www.ordineavvocatimilano.it/it/news/la-carta-dei-principi-per-un-uso-consapevole-dei-sistemi-di-intelligenza-artificiale-in-ambito-forense/p100-n2587>

¹² Le ricerche sull’impatto dell’IA Generativa nel mondo del lavoro sono oramai sterminate. Si menziona in questa sede, per l’autorevolezza della fonte, quello della Banca d’Italia, *An assessment of occupational exposure to artificial intelligence in Italy*, pubblicato a ottobre 2024. Tra gli altri, cfr. anche World Economic Forum, *Jobs of Tomorrow: LLM and Jobs*, Novembre 2023.

2. Dove LLM e Giustizia si incontrano: linguaggio e digitalizzazione

Abstract: I LLM rivelano un'affinità profonda con il diritto grazie alla loro capacità di elaborare il linguaggio naturale, elemento fondante della giustizia. Oltre a ottimizzare l'accesso e l'interpretazione delle fonti normative e supportare l'attività decisionale, gli LLM si integrano nell'ecosistema digitale del diritto contemporaneo, dove ogni fase del processo è ormai tecnologizzata. Tuttavia, il linguaggio giuridico, intrinsecamente complesso e ambiguo, rappresenta una sfida per la piena applicazione degli LLM. Con la loro introduzione, si prospettano trasformazioni radicali, dal supporto nella redazione degli atti alla ricerca giurisprudenziale avanzata, che potrebbero ridefinire il lavoro giuridico nel contesto digitale odierno.

Secondo lo storico Yuval Noah Harari, *“l'IA è la prima tecnologia della storia in grado di prendere decisioni e creare nuove idee da sola. Tutte le precedenti invenzioni umane hanno dato il potere all'uomo perché, per quanto potente fosse il nuovo strumento, la scelta di come e quando usarlo rimaneva nelle nostre maniche”*¹³. Questa definizione centrata sull'abilità decisionale, calata nel contesto dell'attività giurisdizionale, già di per sé genera un inevitabile disagio emotivo. Tuttavia, anche con uno sguardo più ampio, l'affinità elettiva tra giustizia e LLM sembra fondarsi almeno su due fattori: il linguaggio giuridico e la completa digitalizzazione di tutte le fasi del processo, resa possibile dal PAT introdotto dal 1 gennaio 2017.

Sotto il primo profilo, il legame profondo tra diritto e linguaggio non può essere banalizzato, con l'ovvia obiezione che una cosa è il linguaggio, un'altra il ragionamento giuridico. Se infatti il diritto *“è un fenomeno linguistico”* e, almeno *“il diritto moderno – è un discorso, il discorso del ‘legislatore’”*¹⁴, che si applica ai casi concreti mediante le argomentazioni giuridiche, linguaggio e ragionamento devono ritenersi profondamente connessi. Attraverso il linguaggio, sia tecnico-giuridico che descrittivo, la vita reale, di per sé imprevedibile anche per il più avveduto legislatore, entra nelle aule giudiziarie: casi concreti, sempre diversi tra loro, *“raccontati”* in documenti ed atti – anche fotografici - oramai quasi tutti nativi digitali nel procedimento amministrativo; qualificati d'ufficio dal giudice e sussunti nelle fattispecie generali e astratte, a loro volta contenute in formulazioni linguistiche non sempre dettagliate, visto il valore fondamentale dei *“principi”* nel diritto amministrativo¹⁵. Mediante il linguaggio naturale, viene elaborata la decisione in camera di consiglio ed espressa infine la motivazione della decisione ai sensi dell'art.111 della Costituzione, la quale, come è stato acutamente osservato *“non è solamente collegata al modo come il giudice interpreta l'ordinamento nel caso di specie, ma anche al modo come in concreto viene redatta”*¹⁶.

Tanto è centrale e pervasiva oggi l'attenzione al linguaggio, nel contesto della giustizia, che la norma secondo cui gli atti, non solo del giudice, ma di tutte le parti del processo, devono essere

¹³ Y.N. Harari, Nexus. Breve storia delle reti di informazioni dall'età della pietra all'IA, Bompiani, 2024, p. 20 *“Coltelli e bombe non decidono da soli chi uccidere. Sono strumenti ottusi, privi dell'intelligenza necessaria per elaborare le informazioni e prendere decisioni autonome. Al contrario l'IA può elaborare le informazioni da sola e quindi sostituire l'uomo nel processo decisionale. L'IA non è uno strumento, è un agente”*.

¹⁴ R. Guastini, *Il diritto come linguaggio*, Giappichelli, 2006, p. 7; N. Irti, *I cancelli delle parole*, in *Un diritto incalcolabile*, Giappichelli, Torino, 2016, p. 74 *“il diritto ci giunge per il tramite di parole e appieno si identifica con il discorso che uomini rivolgono imperativamente ad altri uomini”*.

¹⁵ Sui principi che anche se di rango costituzionale, in ossequio al principio di legalità, hanno comunque fondamento normativo, cfr. da ultimo, C. Volpe, *Principi e clausole generali nel diritto amministrativo*, in [www.Giustizia - amministrativa.it](http://www.Giustizia-amministrativa.it). 2024

¹⁶ G. P. Cirillo, *Dovere di motivazione e sinteticità degli atti*, in www-giustizia-amministrativa.it.

redatti “*in maniera chiara e sintetica*” (art. 3 c.p.a.)¹⁷, inizialmente introdotta solo nel c.p.a., è diventata una norma processuale di principio (cfr. art. 121 c.p.c.).

La principale obiezione all’inevitabile incontro tra LLM e linguaggio giuridico è la seguente: il linguaggio del diritto è intrinsecamente complesso e difficilmente elaborabile da un computer; è ambiguo, carico di metafore¹⁸, ricco di clausole generali che, negli ordinamenti di diritto positivo, rappresentano “*valvole di apertura volte a favorire l’adeguamento delle formule legislative rispetto al tumultuoso evolversi dei fatti sociali*”¹⁹.

E’ un’obiezione ineccepibile, che però, da un lato è declinabile in ogni dominio a prevalenza linguistica, poiché “*tutti i predicati (...) hanno un riferimento indubbio o aperto (open textured), sicché sono affetti da vaghezza estensionale*”²⁰; e, dall’altro, non considera la rapidità con cui l’evoluzione tecnologica sta trasformando il panorama attuale, lasciando persino gli scienziati sorpresi dalla velocità di questi progressi. In effetti, l’ambiguità linguistica e la difficoltà di comprendere il *contesto*, in cui le parole sono utilizzate nelle conversazioni umane, sono state storicamente tra gli ostacoli principali ai progressi nel campo del *Natural Language Processing*, fino all’avvento dei **Transformer** (cui si riferisce anche la “T” di GPT che sta per “*Generative Pre-trained Transformer*”), i quali si fondano sull’idea chiave di *pesare* l’importanza di ogni parola rispetto alle altre, collocandola in relazione alle altre in una frase o in un intero testo, senza dover analizzare le parole in sequenza come facevano i modelli precedenti, consentendo di cogliere i reciproci nessi semantici anche in testi lunghi e complessi.

L’impatto potenziale dei LLM nell’attività giurisdizionale, trova però, nella giustizia amministrativa, un altro punto di attrazione. Oggi, grazie al PAT, ogni segmento del flusso di lavoro di giudici, avvocati, funzionari è digitale: dal deposito degli atti e documenti, alla loro consultazione, dalla ricerca delle fonti giuridiche e dei precedenti giudiziari, fino alla redazione e pubblicazione delle sentenze. Perfino le udienze, seppure solo quelle di smaltimento, si svolgono in modalità telematica, attraverso la piattaforma Microsoft Teams (alla quale già oggi, per la verbalizzazione, le sintesi delle discussioni o la trascrizione per iscritto, sarebbe applicabile un LLM, come quello che funge da base a Copilot). Un mondo interamente fatto di linguaggio naturale in “*salsa bit*”.

L’ecosistema digitale e linguistico in cui opera la giustizia amministrativa lascia dunque intravedere il potenziale effetto trasformativo dei LLM, come peraltro prefigurato da scienziati di fama mondiale, in relazione anche ad altri sistemi giuridici²¹. Tuttavia, le applicazioni concrete di queste tecnologie restano, per ora, in larga parte ipotetiche, al centro di discussioni dei vertici politici, in cui, condivisibilmente, emergono le esigenze di sicurezza, di qualità dei

¹⁷ L’attenzione legislativa al tema non cala. Da ultimo, con l’art. 1 comma 813 della legge finanziaria 2025 (legge n. 207/2024), all’art. 13 ter comma 5 delle disposizioni di attuazione al c.p.a. è stata anche introdotta una disposizione sanzionatoria pecuniaria, sia pure facoltativa, irrogabile alla parte che in qualsiasi atto del processo superi i limiti dimensionali, senza preventiva autorizzazione; cfr. Cons. Stato, Sez. III, 17 gennaio 2024, n. 352, ordinanza di rimessione all’Adunanza Plenaria sulla questione delle conseguenze della violazione del principio di sinteticità ex art. 3 c.p.a.

¹⁸ F. Galgano, *Le insidie del linguaggio giuridico. Saggio sulle metafore nel diritto*, Il Mulino, 2010.

¹⁹ A. Trabucchi, *Istituzioni di diritto civile*, Wolters Kluwer, ried. 2015, p. 64.

²⁰ R. Guastini, *Prima lezione sull’interpretazione*, Mucchi Editore, 2019, p. 26.

²¹ J Kaplan. cit., pag. 81 “*Le Generative Artificial Intelligence avranno un effetto enorme sulla sanità, ma l’impatto sul diritto sarà ancora maggiore. Nel caso della sanità, è relativamente facile prevedere come le GAI integreranno le procedure attuali. Ma probabilmente le GAI sconvolgeranno il modo in cui scriviamo contratti e direttive, stabiliamo arbitrati e perfino come perseguiamo i criminali. Spesso le questioni mediche prevedono il contatto fisico o la necessità di vedere il paziente, mentre il diritto è soprattutto questione di parole. Un’infinità di parole. E i LLM sono proprio macchine efficienti ed esperte nella gestione delle parole*”

dati, di trasparenza, di mantenimento del principio che la decisione finale resti una “*human-driven activity*”²².

Sui principi chiave, vi è una tendenziale convergenza.

Le difficoltà sorgono quando dai principi si scende al livello delle “azioni”, dei possibili casi di uso dell’IA nel campo dell’attività giurisdizionale. E le problematiche più complesse si incontrano quando principi declamati da anni, come quello della “*spiegabilità*”, ora traslati anche in disposizioni normative positive come fonte di diritti soggettivi (art. 86 AI Act), devono essere applicati in concreto e tutelati, in caso di violazione, nel processo²³. Le possibili applicazioni di seguito accennate intendono essere solo spunti di riflessione, poiché ci troviamo alle soglie di una vera e propria Rivoluzione: la valutazione dei relativi costi-benefici, anche in termini organizzativi e formativi, e non solo finanziari, potrà essere pienamente svolta solo attraverso la sperimentazione sul campo, che permetterà di testarne le reali potenzialità, oltre che verificarne la conformità alle normative vigenti e ai principi del Giusto Processo, cardine dello Stato di diritto.

3. “*Correlation is not causation*”: un mito da sfatare o un limite destinato a durare?

Abstract: I LLM funzionano elaborando contenuti basati su correlazioni probabilistiche tra parole, senza una comprensione reale delle relazioni causali. Questo limita la loro applicabilità in ambiti giuridici centrati sulla causalità, come la responsabilità civile e il diritto penale. Tuttavia, nei compiti giurisdizionali che richiedono correlazioni e sintesi, come l’analisi dei fascicoli o l’esame dei documenti, i LLM possono offrire supporto significativo. Con l’evoluzione tecnologica, i modelli più avanzati iniziano a gestire anche meglio nessi causa-effetto, dimostrando il potenziale di questa tecnologia che è in continua espansione.

Come oramai noto, i LLM elaborano e producono contenuti in linguaggio naturale, basandosi su calcoli probabilistici complessi, sostanzialmente impossibili da decifrare *ex post*; di qui la nota metafora della **Black Box**²⁴ che connota il **Deep Learning** su cui si basano i modelli di oggi.

²² Consiglio dell’Unione Europea, vertice tra i Ministri della Giustizia, Conclusioni sull’Uso dell’IA nella Giustizia, , 16 dicembre 2024, n. 16933/24.

²³ Il tema della trasparenza, della spiegabilità, della comprensibilità delle “decisioni algoritmiche” e dei meccanismi di funzionamento dei sistemi di IA caratterizzati, se fondati sul Deep Learning, dalla cd. Black Box, è ampiamente studiato in letteratura e può essere esaminato in questa sede. Cfr., tra gli altri cfr. V. PAPADOULI, *Transparency in Artificial Intelligence: A Legal Perspective*, 30 maggio 2022, *Journal of Ethics and Legal Technologies*, 4, 25-40. Prospettive ottimistiche fanno leva sul campo di ricerca della cd. **Explainable AI (XAI)** che si pone l’obiettivo di disvelare il meccanismo di funzionamento di dei Sistemi di IA fondati sul Deep Learning. Sull’efficacia, e anche sui tempi brevi dello sviluppo della ricerca in tal senso, si registrano però opinioni discordanti, poiché secondo una parte della scienza si tratterebbe di costruire modelli paralleli che “*approssimano*” il comportamento del sistema di IA che si vorrebbe disvelare, senza alcuna garanzia che la spiegazione fornita per il procedimento parallelo corrisponda effettivamente a ciò che è accaduto nei calcoli svolti in nanosecondi dal sistema: cfr. C. RUDIN, *Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead*, in *Nature Machine Intelligence*, 1, 2019, p. 206-2015. In altri termini, “*benché sarebbe facile programmare un computer affinché stampi una lista delle addizioni e moltiplicazioni eseguite da una rete per un dato ingresso, la lista darebbe a noi umani un’idea pari a zero della strada seguita dalla rete per ricavare la risposta. Una lista di un miliardo di operazioni non è una spiegazione che un essere umano può capire*” M. MITCHELL, *L’intelligenza artificiale. Una guida per essere umani pensanti*, Maverick, 2022, p. 105.

²⁴ La metafora è stata resa celebre da F. PASQUALE, *The black box Society: the secret algorithms that control money and information*, Harvard University Press, 2016

In sostanza, i LLM sono capaci di produrre parole, frasi e interi testi linguistici, predicendo, sulla base dei *pattern* individuati nelle moli di dati di addestramento, quale sarebbe “*parola dopo parola*” quella più adeguata al contesto di riferimento (id est, alla domanda formulata inizialmente dall’interlocutore)²⁵. La predizione avviene a velocità incomparabili con quelle della lettura della mente umana (un essere umano può leggere in media 250 parole al minuto; i LLM oggi più diffusi, milioni), e con efficacia ricorsiva, nel senso che, ad ogni parola aggiunta, il meccanismo si riproduce, fino a completare il testo richiesto. Il valore semantico delle parole è desunto pertanto dal *contesto* linguistico in cui esse sono collocate (secondo il principio formulato dal linguista della semantica distributiva John Firth “*You shall know a word by the company it keeps*”). Ma i LLM non hanno una comprensione della realtà, che è certamente cosa diversa dai dati, né dei nessi causali sottostanti alle informazioni che elaborano. Operando su base puramente statistica, sono perciò notoriamente definiti come dei “*pappagalli stocastici*”²⁶, bravissimi a dire “*cosa*”, ma non “*perché*”.

Le caratteristiche appena menzionate fanno rifuggire i giuristi. Il fenomeno giuridico, infatti, nell’offrire tutela ad interessi ritenuti meritevoli, non è riducibile ad algoritmi linguistici, per quanto complessi, “*perché le forme giuridiche non possono correre il rischio – esiziale – di separarsi da valori e interessi diffusi e condivisi*”²⁷. Inoltre, almeno nei settori del diritto penale e della responsabilità per danni anche della pubblica amministrazione, è centrale il principio di causalità tra eventi e conseguenze. Il più grande limite dell’IA applicata al mondo giuridico viene così racchiuso nel mantra “*correlation is not causation*”.

Tuttavia, a ben vedere, questa affermazione assume talvolta i contorni di un mito.

In primo luogo, le correlazioni – se accurate e verificabili – possono rivelarsi strumenti preziosi per il lavoro del giudice, non solo nell’analisi degli atti e dei documenti probatori, ma anche nella ricerca giuridica e nell’assistenza alla redazione dei provvedimenti. Il valore dell’informazione, infatti, non dipende esclusivamente dalla sua natura causale o consequenziale, ma anche dalla sua capacità di offrire nuovi punti di osservazione e connessioni utili al ragionamento giuridico, ampliando la base di conoscenza.

In secondo luogo, il concetto stesso di causalità è destinato a essere ridefinito con l’evoluzione degli strumenti di IA. I modelli di ultima generazione, ben diversi da quelli degli esordi, non si limitano più a individuare mere correlazioni statistiche, ma sono in grado di cogliere schemi complessi, individuare relazioni latenti tra dati e suggerire inferenze più sofisticate²⁸. Il punto

²⁵ M. Suleyman, *L’onda che verrà. Intelligenza artificiale e potere nel XXI secolo: “Gli LLM sfruttano la sequenzialità dei dati linguistici. Ogni unità di informazione è in qualche modo correlata ai dati che la precedono nella serie. Il modello legge un numero elevatissimo di frasi, impara una rappresentazione astratta delle informazioni contenute in esse e quindi genera una previsione su cosa dovrebbe seguire. La sfida consiste nel progettare un algoritmo che “sappia dove cercare” i segnali di una data frase. Quali sono le parole chiave, gli elementi salienti di una frase in modo che si correlano tra loro? E’ quanto nell’IA viene definito “attenzione”. Il meccanismo dell’attenzione è alla base del cd. Transformer ed è stato inventato da alcuni ricercatori di Google nel 2017, ed è diventato oggetto di studi grazie al paper da loro pubblicato dal titolo evocativo della più nota canzone dei Beatles “Attention is all you need”, uno degli articoli più ricercati sul web dopo l’apparizione sulla scena di ChatGPT.*

²⁶ La locuzione, oramai di uso comune, è stata coniata da E. M. Bender, T. Gebru, A. McMillan-Major, S. Shmitchell, *On the dangers of stochastic parrots: can Language Models be too Big?*, 2021, <https://doi.org/10.1145/3442188.3445922>

²⁷ P. Grossi, *L’invenzione del diritto*, Editori Laterza, 2017, 113

²⁸ Secondo gli osservato, peraltro, i LLM più avanzati hanno dimostrato infatti di comprendere, nell’ambito di testi linguistici anche articolati, le relazioni di causa-effetto, riconoscendo ad esempio nei contratti la clausola che contiene una condizione rispetto all’evento-conseguenza. Questo progresso nelle abilità è stato osservato dai ricercatori anche in relazione ad altre abilità, peraltro fondamentali nel metodo giuridico, come il ragionamento sociale, cioè la capacità di comprendere e simulare scenari in cui ci sono motivazioni e prospettive diverse tra più individui e, per entrambe tali “abilità”, sono stati elaborati benchmark specifici, proprio con lo scopo di valutare la capacità dei più avanzati LLM. Il

centrale, quindi, non è tanto il superamento della distinzione tra correlazione e causalità, ma piuttosto la comprensione di come l'IA possa integrarsi nel lavoro giuridico senza sostituire la funzione critica del giurista e la natura umana della decisione, idonea ad incidere sulle sfere giuridiche di cittadini, imprese, amministrazioni. La sfida che occorrerebbe cogliere è come valorizzare queste tecnologie, oramai già entrate nella vita quotidiana, come strumenti di supporto anche all'attività lavorativa, senza rinunciare al ruolo umano di interpretazione e garanzia della correttezza giuridica ed avendo piena consapevolezza dei meccanismi di funzionamento, dei limiti, ma anche delle possibili misure di mitigazione dei rischi.

4. La Giustizia nell'AI Act e le vie di fuga

*Abstract: Il Regolamento UE 2024/1689 (AI Act) classifica i sistemi di Intelligenza Artificiale destinati alla Giustizia come "ad alto rischio", sottoponendoli a requisiti stringenti. Tuttavia, l'applicazione dell'IA alla Giustizia solleva interrogativi sia normativi, prevedendo eccezioni ancora da definire; che operativi, specialmente con l'uso dei LLM. Il regolamento si riferisce solo all'utilizzo giurisdizionale dell'IA, evidenziando un possibile divario digitale tra magistrati e le professioni legali, che abbracciano con maggiore rapidità le tecnologie Legal Tech. La supervisione umana rimane centrale, ma richiede nuove competenze, come il **prompt engineering** e la valutazione critica dei risultati generati dall'IA.*

La Giustizia è ritenuta dal legislatore dell'Unione Europea un settore "sensibile", poiché rientra nell'Allegato III del Regolamento UE 2024/1689 (AI Act) rubricato "Sistemi di IA ad alto rischio di cui all'articolo 6 paragrafo 2"²⁹. Ciò significa che ogni **Sistema di IA** che sia progettato e sviluppato per tale contesto –come negli altri elencati nell'Allegato III - anche se basato su un LLM, comporta l'applicazione della congerie di norme, requisiti, sanzioni previsti dal Capo III del Regolamento "Sistemi di IA ad alto rischio", che si apre con l'articolo 6.

Peraltro, per quanto non possa essere approfondito in questa sede il profilo regolatorio, se un sistema di IA dedicato alla Giustizia si fonda su un LLM, ad esso andrebbero applicati due compendi normativi contenuti nell'AI Act: sia le disposizioni del Capo III, citato, sia quelle del Capo V – inserito nell'impianto dell'IA Act durante la fase parlamentare per effetto dello tsunami dell'IA Generativa – che riguarda i "modelli general purpose" e, quindi, anche i LLM³⁰. Un bel grattacapo, visto l'elevato sforzo interpretativo richiesto dal Regolamento che

report annuale di Stanford del 2024, ormai risalente ad aprile dello scorso anno, dà conto dei risultati raggiunti da LLM come GPT 4 sul ragionamento sociale, riportando i test effettuati con BigToM (*Theory of Mind*).

²⁹ Secondo il punto 8 dell'allegato III, sono considerati ad alto rischio i sistemi di IA destinati a essere utilizzati "da un'autorità giudiziaria o per suo conto, per assisterla nella ricerca e nell'interpretazione dei fatti e del diritto, o nell'applicazione della legge a una serie concreta di fatti o in modo analogo nella risoluzione alternativa delle controversie."

³⁰ Come è noto, il tema dell'AI applicato alla giustizia è trattato anche nel DDL approvato dal Consiglio dei Ministri a maggio 2024 e ora al Senato in senso molto più restrittivo e le divergenze tra il Regolamento e la proposta normativa italiana non sono limitate solo a questo ambito, con conseguenti delicate questioni di compatibilità. Giova osservare che l'AI Act è adottato ai sensi dell'art. 288 comma 2 del TFUE, sulla base dell'art. 114 quale norma legittimante; che dal momento in cui l'Unione emana un regolamento in un settore, "gli Stati membri sono tenuti ad astenersi da qualsiasi provvedimento che deroghi a tali regolamenti o ne pregiudichi l'efficacia" (Corte giust., 2 febbraio 1977, 50/76); che il disegno di legge italiano non si occupa solo di profili delegati agli Stati membri dall'AI Act, come quelli sulla governance o sulle sanzioni, ma contiene norme che hanno il medesimo ambito oggettivo delle norme europee e vi si sovrappongono, come quelle sulle definizioni, o sui diversi settori di applicazione; che la Commissione europea ha trasmesso all'Italia, il 5 novembre 2024, un parere circostanziato (C(2024) 7814) sul DDL in cui, tra l'altro, in riferimento all'articolo 14, che limita l'utilizzo dei sistemi di IA nell'attività giudiziaria solo "per l'organizzazione e semplificazione del lavoro giudiziario e per la ricerca giurisprudenziale e dottrinale", invitando ad allineare tale norma, all'articolo 6, paragrafo 3, del regolamento sull'IA. In relazione al Capo V, però, vista la recente emersione dei modelli di IA Generativa, l'AI Act rimanda ad una forma di co-regolazione, prevedendo l'adozione di un "Codice di condotta", la cui prima bozza è stata pubblicata a novembre 2024. Cfr. <https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practice-published-written-i>

in parte è entrato in vigore il 2 febbraio 2025 (per i sistemi di IA vietati e per l'obbligo di "alfabetizzazione" di cui all'art. 4).

La definizione normativa della "Giustizia" contenuta nell'Allegato III conduce però a due osservazioni preliminari, quanto all'ambito applicativo.

Va innanzi tutto considerato che essa è centrata solo sulla figura dell'autorità giurisdizionale e di quella, assimilata, dell'arbitro, quindi la giustizia è correttamente intesa come sinonimo di "attività giurisdizionale". Non vi è cioè nessun cenno ai possibili usi di sistemi di IA nell'esercizio della professione forense (mentre un ampio spazio nel Regolamento è riservato all'altra parte del processo amministrativo, la pubblica amministrazione, con il conseguente rilevante tema del sindacato giudiziale sui vizi della decisione algoritmica³¹), verosimilmente sul presupposto che, per quanto "*professione regolamentata*" si tratti comunque dell'esercizio di "*prestazioni di servizi professionali*". Senonché, il servizio erogato dalla giustizia amministrativa, come delineato dal codice del processo amministrativo, è il frutto di tutte le parti coinvolte nel processo, le quali collaborano per la realizzazione del principio del giusto processo, ai sensi dell'art. 2 comma 2 c.p.a.

Il tema, può apparire fuori fuoco, poiché è evidente che l'impatto sulla vita dei cittadini deriva dalle decisioni giurisdizionali, ma potrebbe rivelarsi, nel tempo, foriero di divergenze sia di preparazione tecnica che culturali tra giudici ed avvocati, destinate a incrinare l'efficienza di un sistema che, dal punto di vista del cittadino, è funzionalmente unitario.

Il rapporto del 2024 sul **Legal Tech** in Italia pubblicato dal Gruppo Sole 24 Ore³², evidenzia infatti che questo settore è in rapidissima espansione e che l'effetto trasformativo dell'IA sulle professioni legali è già avviata, ridisegnando nel profondo le competenze richieste agli avvocati: abilità tradizionali come la rapidità di scrittura e la capacità di sintesi sarebbero destinate a perdere centralità, lasciando il posto a competenze digitali e a *soft skills* legate alla strategia e alla capacità di visione d'insieme.

La focalizzazione del "*rischio alto*" da IA solo sull'attività imparziale del giudice/arbitro contenuta nell'AI Act sembra in sostanza dar luogo ad uno scollamento delle norme sia dalla realtà processuale, che dalla realtà economica e rappresentare l'abbrivio di una nuova forma di *digital divide* tra professionalità diverse nell'ambito dell'unico ecosistema "Giustizia", destinata inevitabilmente ad ampliarsi, se si considera che l'utilizzo delle tecnologie di IA, come ausilio nello svolgimento del proprio lavoro richiede, non solo formazione teorica, ma soprattutto esperienza applicativa svolta sul campo.

Il secondo aspetto da sottolineare riguarda il rapporto, interno all'AI Act, tra regola ed eccezioni. Il paragrafo 3 dell'articolo 6 del Capo III prevede che, per essere qualificato alto, il rischio deve essere "*significativo per i diritti fondamentali e i valori dell'UE*" e che il sistema di IA deve essere in grado di "*influenzare materialmente il risultato decisionale*". Lo stesso paragrafo prevede però anche ipotesi esemplificative: non è, ad esempio, ad *alto rischio* il

³¹ A. Corrado, *Il sindacato giurisdizionale sull'attività amministrativa algoritmica*, in *Il sindacato giurisdizionale sull'attività amministrativa* (a cura di) A. Cassatella, R. Chieppa, A. Moliterni, La Tribuna, 2024, pag. 481 e ss.

³² *Il Legal Tech in Italia. Dati, soluzioni e opinioni*, Gruppo Sole24Ore., cfr, in particolare, G. Lega e E. Raffiotta, *IA nel futuro delle professioni legali*, Rapporto cit., pag. 8.

sistema di IA che esegua compiti procedurali limitati³³, come la gestione dei fascicoli, l'indicizzazione, la ricerca, la traduzione di documenti; o che sia diretto a “migliorare il risultato di un'attività umana già completata”; o che sia utilizzato per individuare schemi decisionali o deviazioni da decisioni precedenti, senza sostituire o influenzare in modo determinante il giudizio umano³⁴.

L'utilizzo di sistemi di IA, anche basati su LLM, per “migliorare” il risultato dell'attività svolta dall'essere umano merita una particolare attenzione, nel contesto dell'IA Generativa. Essa presuppone una demarcazione tra attività riferibile all'intelligenza “umana” e attività invece imputabile a un'applicazione di intelligenza “artificiale”. Ora, questa distinzione, tra apporto dell'uomo e contributo “ausiliario” della macchina, assume toni molto sfumati nella pratica operativa dei LLM. Ciò che connota tali modelli è infatti l'interazione circolare e continua tra l'interlocutore umano e il “computer”, rappresentando vividamente il principio **Human-in-the-loop**.

Nei sistemi di IA che si fondano sui modelli linguistici, come ChatGPT, per citare il più noto, l'essere umano fornisce un input iniziale (*prompt*), ma, a differenza del tradizionale quesito inserito nei motori di ricerca, la risposta iniziale fornita dal sistema è solo l'avvio di una relazione dinamica che si dipana in modo interattivo. Il processo dialogico si sviluppa attraverso una serie di scambi che partono dall'input iniziale, ma poi si articolano in istruzioni dettagliate, approfondimenti su questa o quella parte, richieste di uno specifico stile, indicazioni sulle fonti da utilizzare, limitazioni nel numero di parole, divisione in sezioni, richieste di sintesi³⁵, comparazioni tra più argomenti/informazioni. Il risultato finale, ossia quello che è ritenuto soddisfacente rispetto alla propria esigenza conoscitiva dall'utente, è il frutto di questa interazione, di una sorta di “lavoro di squadra” in cui non sono però identificabili i singoli contributi delle diverse “intelligenze”.

Se, nella Giustizia deve pertanto restare ferma la garanzia che il giudice mantenga il ruolo di *dominus* della decisione finale, in linea con i principi costituzionali vigenti³⁶ e che la “supervisione umana” non si riduca a un mero simulacro, è imprescindibile un approccio

³³ Secondo il Considerando 53, i compiti procedurali limitati includono funzionalità come l'indicizzazione, l'elaborazione testuale o vocale e il collegamento con altre fonti di dati, tutte attività di particolare rilevanza nel processo amministrativo telematico. Tuttavia si tratta di attività che dovrebbero anche tenersi distinte da quelle “puramente accessorie” come l'anonimizzazione o la pseudonimizzazione delle decisioni, documenti o atti giudiziari, o i compiti amministrativi che, già a prescindere dalle eccezioni, comunque non rientrerebbero nella categoria dei sistemi ad alto rischio applicati alla Giustizia (Considerando 61, ultimo periodo). Rientrano in questo ambito, i progetti già in corso nella Giustizia amministrativa, su cui cfr. infra nota 48.

³⁴ L'ambito di applicazione delle eccezioni di cui al paragrafo 3 citato sarà meglio definito quando, entro il 2 febbraio 2026, saranno adottati gli “orientamenti” della Commissione sul punto, come previsto dall'art. 6 par 5. Ma questa prefigurata netta demarcazione tra “attività umana” e apporto dell'IA appare di per sé ambigua, se i sistemi di IA in questione si basano su LLM. Anche in questo caso, un work in progress.

³⁵ Per alcuni ricercatori dell'Università Cattolica del Sacro Cuore, questa relazione iterativa tra l'uomo e l'IA rappresenta un vero e proprio cambio di paradigma, il cuore di una vera e propria rivoluzione mentale. L'uso dei LLM non si limita a fornire strumenti che velocizzano processi preesistenti, ma modifica profondamente il modo di pensare dell'utente stesso. I modelli linguistici si integrerebbero così nei processi decisionali umani, combinando i due sistemi di pensiero elaborati dall'economista Premio Nobel Daniel Kahneman: il “Sistema 1” (decisioni intuitive, rapide e automatiche) e il “Sistema 2” (decisioni razionali, lente e analitiche), configurandosi come un nuovo “Sistema 0”. C. Galletti, *Sistema 0, l'intelligenza artificiale sta già cambiando il cervello umano: cosa ha scoperto una ricerca italiana*, in https://www.corriere.it/tecnologia/24_ottobre_26. La ricerca è stata anticipata da alcune riflessioni di uno dei coordinatori del team, M. Chiriatti, *Incoscienza artificiale. Come fanno le macchine a prevedere per noi*, Luiss, 2021.

³⁶ Per una ricostruzione scientifica della natura “eminente umana” della giurisdizione nell'alveo costituzionale, G. Gallone, *Riserva di umanità, intelligenza artificiale e funzione giurisdizionale alla luce dell'AI Act. Considerazioni (e qualche proposta) attorno al processo amministrativo che verrà*, in www.giustizia-amministrativa.it. Da ultimo, M. Interlandi, *Large Language Model e sistema di giustizia amministrativa. Prime riflessioni sul ruolo del giudice amministrativo*, in www.giustizia-amministrativa.it.

consapevole e pragmatico che non guardi all'IA come ad un *unicum*, ma alla specifica tecnologia con la quale ci si sta interfacciando e al suo concreto modo di operare. Non basta invocare principi astratti, che potrebbero solo rinforzare l'illusione del controllo; occorre misurarsi direttamente con il funzionamento concreto di questi strumenti, comprenderne le reali implicazioni operative, sperimentarne le potenzialità e i limiti nei casi applicativi, prima di farli entrare nell'attività quotidiana.

5. I Rischi dei LLM applicati alla Giustizia e le possibili misure di mitigazione

Abstract: L'applicazione dei LLM alla giustizia presenta tre rischi principali: automation bias, allucinazioni e appiattimento sulla cultura giuridica del common law. L'automation bias riflette una fiducia eccessiva nei sistemi tecnologici, che può compromettere l'indipendenza decisionale umana, mentre le allucinazioni riguardano la generazione di risposte non supportate da dati reali. Entrambi i rischi richiedono misure di mitigazione, come formazione specifica, tecniche di Retrieval-Augmented Generation e il Prompt Engineering. Il rischio di appiattimento sulla cultura del common law, invece, deriva dall'affinità dei LLM con i metodi induttivi tipici di quei sistemi, in contrasto con la logica deduttiva dei sistemi di Civil Law. Mitigare questi rischi implica integrare criteri interpretativi normativi nel fine-tuning dei modelli e promuovere consapevolezza sull'importanza della specificità giuridica degli ordinamenti di civil law.

I rischi associati all'IA sono da anni al centro dell'agenda dei regolatori globali. Il *Considerando 1* dell'AI Act prevede che l'obiettivo del Regolamento, adottato ai sensi dell'art. 114 TFUE, è quello di promuovere “*un'intelligenza artificiale (IA) antropocentrica e affidabile*”, garantendo però un elevato livello di protezione della salute, della sicurezza e dei diritti fondamentali sanciti dalla Carta dei diritti fondamentali dell'Unione europea, compresi i principi di democrazia, Stato di diritto e protezione dell'ambiente (questi ultimi tre, entrati nel testo all'esito della fase parlamentare). Nell'ambito della funzione giurisdizionale, i rischi normalmente associati all'IA (perdita del controllo, *bias* amplificati, violazione della privacy, disinformazione, insondabilità) acquisiscono una dimensione concretamente percepibile, perché le decisioni giurisdizionali incidono in modo diretto e profondo sulla vita dei destinatari finali, rendendo cruciale l'affidabilità dei sistemi di IA e la loro governabilità da parte dell'uomo. Quelli che seguono intendono essere esempi di un approccio costruttivo, che non si limiti a individuare i problemi posti dall'innovazione tecnologica, ma che, anche attraverso la stessa tecnologia, proponga possibili soluzioni.

5.1. Il rischio di automation bias

L' *automation bias* consiste nella tendenza a riporre una fiducia eccessiva nei suggerimenti forniti da un LLM, ignorando o sottovalutando segnali divergenti, provenienti da altre fonti. E' un fenomeno ben noto in altri settori a forte automazione, come l'aviazione civile³⁷ e, nel contesto della Giustizia, rischia di rendere astratto il principio, più volte richiamato, secondo

³⁷ K. Mosier, L. Skitka, M. Dunbar, L. McDonnell, *Aircrews and Automation Bias: The Advantages of Teamwork?*, in *The international journal of aviation psychology*, 2001, 11, 1-14, nel quale si evidenzia come la disponibilità di strumenti automatizzati alimenti la tendenza umana generale a scegliere il percorso di minore sforzo cognitivo, affidandosi ciecamente alla macchina. Ha segnato una svolta in questi studi il disastro aereo dell'Air France 447, il 1 giugno del 2009, precipitato nell'oceano Atlantico, causando la morte di 228 persone, tra passeggeri ed equipaggio. Le investigazioni successive hanno poi rivelato che l'errore tecnico sarebbe stato trascurabile, se non vi fossero stati, in pochi minuti, diversi errori umani proprio relativi anche alle interazioni tra i piloti e i sistemi automatici di pilotaggio, che avevano dato segnali di allarme ignorati.

cui “l’utilizzo di strumenti di IA può fornire sostegno al potere decisionale dei giudici o all’indipendenza del potere giudiziario, ma non dovrebbe sostituirlo: il processo decisionale finale deve rimanere un’attività a guida umana” (Considerando 61 dell’AI Act).

I sintomi dell’*automation bias* possono manifestarsi come errori di omissione, se non sono valorizzati elementi rilevanti in assenza di un suggerimento diretto da parte del sistema di IA, pur disponendo di fonti diverse; come errori di commissione, se le proposte del LLM sono accettate senza una verifica critica ed autonoma da parte di colui cui è imputabile la decisione; mentre le cause del fenomeno, studiate nelle scienze comportamentali, includono l’avarizia cognitiva, consistente nella tendenza umana a ridurre lo sforzo decisionale. Il bias può fondarsi sulla *complacency*, ossia nella fiducia eccessiva nelle tecnologie, specie dopo una prima fase di “rodaggio” e maggiore prudenza; sul *pro-innovation bias*, consistente nell’ottimismo acritico verso l’innovazione tecnologica; o, più in generale, sulla tendenza degli esseri umani alla “*avarizia cognitiva*”.

L’AI Act accenna al rischio di *automation bias* nell’art. 14 “Sorveglianza umana” – (disposizione che si applicherebbe però solo per i Sistemi di IA ad alto rischio, essendo inserito nel citato Capo III). In particolare, il paragrafo 4 lettera b) dell’art. 14 stabilisce che le persone, preposte alla sorveglianza dei sistemi di IA, debbano restare consapevoli «della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull’output prodotto da un sistema di IA ad alto rischio (“*automation bias*”), in particolare per i sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche». L’obbligo di tener conto di questa tendenza è da apprezzare positivamente, perché in linea con la spinta dell’Unione Europea a considerare approcci innovativi come quelli offerti proprio dalle scienze comportamentali, per la elaborazione di “*regole a prova di futuro*”³⁸.

Ma quali potrebbero essere le misure concrete per evitare che la validazione della “proposta” di un LLM diventi meramente formale e si trasformi in una sorta di nuova “burocrazia digitale” dove alla paura della firma si sostituisce la paura di decidere in maniera divergente dalla proposta algoritmica, dovendo darne conto? Probabilmente, la strategia vincente è quella combinata: progettazione tecnica degli LLM con *alert* specifici ed interfaccia che garantiscano la ponderazione della decisione; formazione mirata di chi si avvale del LLM, diretta a sviluppare competenze critiche e le abilità di riconoscere e gestire eventuali errori sistemici, anche mediante simulazioni pratiche³⁹; esplicabilità del metodo di “ragionamento” seguito dal sistema. Ma, occorrerebbe anche aprire una riflessione sulle misure organizzative e processuali, ad esempio imponendo sempre e comunque un sistema di doppia verifica e validazione collegiale (escludendo decisioni monocratiche), qualora a supporto dell’attività giurisdizionale ci si avvalga di un LLM. Quest’ultima potrebbe sembrare una misura bizzarra. Eppure è di fatto già prevista nell’AI Act, con riferimento ai sistemi di riconoscimento biometrico in tempo reale e in spazi aperti al pubblico e nulla esclude che il legislatore nazionale possa estenderne l’applicazione oltre i casi minimi obbligatori. Nell’ipotesi –che costituisce un’eccezione al

³⁸ Nella Comunicazione della Commissione UE “*Legiferare meglio: unire le forze per produrre leggi migliori*” del 29 aprile 2021 si prende atto dell’effetto dirompente che i cambiamenti dell’era digitale hanno anche sugli approcci e sugli strumenti a disposizione di regolatori, e nel *Better Regulation Toolbox* del 25 novembre 2021 indica l’uso delle scienze comportamentali e la *regulatory sandbox* come metodi emergenti utilizzabili per elaborare regole future-proof.

³⁹ S. Alon-Barkat, M. Busuioc-Human, *AI Interactions in Public Sector Decision Making: “Automation Bias” and “Selective Adherence” to Algorithmic Advice*, in *Journal of Public Administration Research and Theory*, 2023, 33, 153 e ss

divieto generale— in cui tali sistemi particolarmente invasivi siano ammessi, l’art. 14 paragrafo 5⁴⁰, prevede infatti che la sorveglianza umana debba essere affidata “*a due persone fisiche separatamente*”, sul presupposto che il controllo affidato ad un solo essere umano non dia sufficiente garanzia (e proprio sulla falsariga del pilotaggio civile aereo, in cui il fenomeno dell’automation bias è ampiamente studiato).

5.2. Il rischio di allucinazioni

Il rischio di *allucinazioni* dei LLM è quello più discusso nel dominio legale⁴¹. Si tratta di un rischio intrinseco al loro funzionamento, trattandosi di modelli addestrati su enormi quantità di dati, estrapolati dal *web*, in cui un articolo scientifico di *Nature* e un post di *Instagram* contengono parole valutate della medesima “credibilità” nella fase di addestramento. Il termine *allucinazione* è una metafora che, nella diffusa antropomorfizzazione della tecnologia, evoca l’idea i LLM sappiano “mentire” o comunque abbiano la tendenza a generare risposte non supportate da dati reali o verificabili. D’altra parte, non si tratta dell’unica metafora che l’IA porta con sé. Tra gli effetti dirompenti di questa tecnologia, emerge anche una nuova nomenclatura, carica di definizioni evocative volte ad “antropomorfizzarla”: lo è in fondo anche la nozione di “apprendimento automatico”, con cui si traduce l’approccio del **Machine Learning**, poiché le macchine non “imparano”, come fanno i bambini, ma riconoscono nei dati correlazioni e poi individuano “con un certo livello di autonomia” soluzioni per raggiungere l’obiettivo assegnato.

Le allucinazioni sono un limite sperimentato da chiunque si sia interfacciato con LLM generalisti, come ChatGPT di OpenAI o *Gemini* di Google, e che, se può essere tollerato, quando si richiede al modello di scrivere poesie, ricette vegane o il prossimo piano per le vacanze, conduce a più che giustificate resistenze, quando il dominio di applicazione, come quello giuridico, impone affidabilità e accuratezza.

Gli scienziati attribuiscono questo limite al cd. “rumore di fondo” dei dati di addestramento, una mole enorme, un’infinità di parole estrapolate online non danno nessuna garanzia di qualità, secondo il noto aforisma “*Garbage In Garbage Out*”. Ma i fattori causali individuati per le allucinazioni sono anche altri: la mancanza di comprensione della realtà cui si riferiscono le informazioni tratte del mondo digitale; la mancanza di aggiornamento nei dati di addestramento rispetto ad un argomento specifico; *prompt* troppo vaghi o, al contrario, troppo specifici, così da generare la tendenza dei LLM ad “assecondare” la domanda in modo acritico, anche quando essa stessa contiene informazioni inesatte (cd. Alignment Bias); in prospettiva, con la quantità di contenuti sintetici in circolazione generati dalla stessa IA, (cd. *model collapse*), persino il

⁴⁰ “per i sistemi di IA ad alto rischio di cui all’allegato III, punto 1, lettera a), le misure di cui al paragrafo 3 del presente articolo sono tali da garantire che il deployer non compia azioni o adotti decisioni sulla base dell’identificazione risultante dal sistema, a meno che tale identificazione non sia stata verificata e confermata separatamente da almeno due persone fisiche dotate della necessaria competenza, formazione e autorità”.

⁴¹ In questo campo, la ricerca ha individuato diverse tipologie di “allucinazioni giuridiche”, enfatizzando molto le risposte errate che tendono ad assecondare le premesse, magari errate o non accurate, contenute nel prompt, amplificandone le inesattezze: cfr. M. Dahl, V. Magesh, M. Suzgun, D. E. Ho, *Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models*, in <https://academic.oup.com/jla/article/16/1/64/7699227>, 26 giugno 2024. Il tema della allucinazioni ha avuto eco mediatica per l’effetto di note vicende giudiziarie, causate più dalla mancanza di professionalità degli avvocati protagonisti che dalla tecnologia stessa. Nel maggio 2023, l’avvocato Steven A. Schwartz pensò di utilizzare ChatGPT per un caso a cui stava lavorando, avente ad oggetto l’azione di risarcimento intentata da un passeggero, Roberto Mata, nei confronti della compagnia aerea Avianca. Per semplificare la ricerca che doveva fare, l’avvocato chiese a ChatGPT un elenco di precedenti giudiziari, poi multa di 5000 dollari, è però diventato famoso, avendo avuto la vicenda risonanza in tutto il mondo. Cfr. <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>.

progressivo degradarsi della qualità dell'output derivante dall'uso di dati generati dall'IA senza validazione. Certo anche questo fenomeno è destinato a diminuire, con l'evoluzione tecnologica. Basti solo considerare che le più recenti versioni dei LLM "generalisti" allegano alla risposta la fonte (seppure come *link* ad una pagina *web*) da cui l'informazione è stata desunta e rielaborata; un primo step, non esaustivo, per saggiarne anche la attendibilità.

Individuate le cause, può apparire più semplice anche intravedere possibili misure di mitigazione. Fondamentale è naturalmente l'addestramento dei LLM su dataset accurati e aggiornati, supervisionati da esperti legali dello specifico settore giuridico di utilizzo. Le tecniche più avanzate in questa direzione suggeriscono il **Retrieval-Augmented Generation** (RAG), combinando così le abilità linguistiche tradizionali dei LLM che sono state consentite dal loro essere "large", con il supporto di archivi "di qualità" selezionati per lo specifico contesto e guidati dagli esperti del settore (magistrati, avvocati, funzionari del processo), in modo tale che il LLM, prima di generare le risposte con le proprie capacità linguistiche, attinga le informazioni da tali fonti affidabili e validate. Il rischio di allucinazione può naturalmente essere mitigato sia con modelli che forniscano, accanto alla risposta, anche la fonte specifica di riferimento (come già oggi avviene, come accennato), sia con una adeguata formazione degli interlocutori sulle tecniche di **Prompt Engineering**, che nel contesto giuridico è stato già declinato come "**Legal Prompting**", ossia sul cruscotto che dovrebbe abilitare alla guida della macchina.

5.3. *Il rischio di omologazione alla cultura giuridica del Common Law*

Un rischio piuttosto trascurato nella letteratura italiana deriva dalla assoluta dominanza, nei *dataset* di addestramento e nella progettazione dei LLM della cultura giuridica di *Common Law*⁴², a sua volta determinata dalla assoluta predominanza, nella geopolitica dell'IA Generativa, di Stati Uniti e UK, oltre che della alla Cina⁴³. Non è casuale che il binomio IA/Giustizia si accompagni frequentemente alla nozione, ancora tutta da definire, di "giustizia

⁴² Per un criterio di distinzione utile nell'alveo dell'IA applicata alla giustizia, cfr. Barberis, *Giustizia predittiva: ausiliare e sostitutiva. Un approccio evolutivo*, in *Milan Law Review*, 2022, 3 "La differenza principale fra le due culture non sta nel carattere rispettivamente obbligatorio o persuasivo del precedente, come dicono i comparatisti. Piuttosto, nel common law la decisione giudiziale può prendersi anche sulla base d'un solo precedente – inclusa la precedente applicazione d'una legge – mentre nel civil law va sempre motivata in base a una legge. Di qui il diverso ruolo dell'interpretazione, nel suo senso tipicamente continentale di attribuzione di significato a testi. L'interpretazione, quasi irrilevante per i precedenti di common law, privi di formulazione canonica, ossia costitutiva della norma stessa, è invece tremendamente sopravvalutata nel civil law. La teoria analitica continentale, specie il realismo "genovese", ritiene decisiva per la decisione giudiziale l'attribuzione di significato ai testi. Come sa qualsiasi giurista pratico, invece, tale attribuzione dipende meno dai testi che dai fatti stessi oggetto di decisione". Naturalmente l'osmosi tra i due approcci culturali è oggetto di approfonditi studi che, quanto al diritto italiano, ad esempio pongono l'accento sul tema del "diritto vivente" oramai entrato "nel vocabolario della Corte costituzionale per dare concretezza, oltre il mero dato positivo, all'oggetto del suo sindacato di costituzionalità", D. Sabatino, *La via italiana al precedente vincolante*, in *Giustizia-amministrativa.it*, 2013; cfr. da ultimo, Corte Cost. 203/2024 "il tendenziale rispetto dei propri precedenti – unitamente alla coerenza dell'interpretazione con il testo delle norme interpretate e alla persuasività delle motivazioni – è per le giurisdizioni superiori, condizione essenziale dell'autorevolezza delle loro decisioni".

⁴³ Il tema geopolitico dell'IA è oggi più che mai sotto i riflettori. Per una accurata e aggiornata panoramica delle relative questioni anche giuridiche, cfr. A. Malaschini, *Le grandi società tecnologiche e le nuove forme della statualità*, in *Rassegna parlamentare*, n. 3/2024. E' di qualche giorno fa l'exploit di DeepSeek, un modello linguistico open source lanciato da una start up cinese che ha causato in poche ore un crollo dei titoli di aziende americane del settore dell'IA: Nvidia, Oracle, Microsoft e Alphabet, la società madre di Google, considerata una delle mosse della guerra fredda tecnologiche che è ormai in atto tra Stati Uniti e Cina per rivendicare il primato mondiale nel campo dell'IA.

*predittiva*⁴⁴, in cui il *focus* è spostato sulla capacità predittiva dei sistemi di IA, basata sulla elaborazione dei “*precedenti giurisprudenziali*”. Vero è però che, in ordinamenti giuridici di *Civil Law*, come quello italiano ed eurounitario⁴⁵, il parametro fondamentale su cui motivare la decisione ex art. 111 Cost. nel caso concreto resta sempre quello normativo, anche se è racchiuso in un principio costituzionale o in un principio generale di diritto amministrativo. Il rischio è pertanto che LLM progettati *negli e per* gli ordinamenti anglosassoni rechino una eccessiva enfasi sul formante giurisprudenziale, trascurando il dato normativo e le tecniche interpretative, anch’esse, almeno nell’ordinamento italiano, disciplinate dal diritto positivo ai sensi degli artt. 12 e ss delle disposizioni preliminari al codice civile e dal primato dell’interpretazione testuale, più volte ribadito anche recentemente dalla giurisprudenza amministrativa⁴⁶ (e nonostante l’origine “pretoria” del diritto amministrativo).

Anche in relazione a questo rischio, misure di mitigazione possono essere suggerite dalla stessa tecnica, come la progettazione di LLM che inglobi tra i parametri di elaborazione delle fonti proprio i criteri interpretativi delle norme⁴⁷ o che, tra i precedenti giudiziari e la letteratura giuridica, dia maggior peso alle fonti nelle quali l’argomentazione tiene conto proprio dei criteri ermeneutici, anche se, come doverosamente accade, volti ad adattare il diritto positivo all’emergere di nuove esigenze ancora non affrontate dal legislatore.

⁴⁴ Sulla difficile delimitazione della definizione di “giustizia predittiva” e sulle plurime questioni che il tema pone, cfr. G. Carloti, *Giustizia predittiva e le fragole con la panna*, in giustizia-amministrativa.it “Per i fini delle mie odierne esigenze espositive, mi attesto, dunque, su una definizione del tutto provvisoria e soltanto descrittiva del fenomeno. Mi riferirò ai programmi di giustizia predittiva come degli algoritmi di apprendimento automatico (machine learning) che utilizzano le tecniche dell’elaborazione del linguaggio naturale (natural language processing) e dell’apprendimento profondo (deep learning) per analizzare, tramite la ricerca testuale e la clusterizzazione, dei megadati (ossia grandi insiemi di dati) - rappresentati da provvedimenti normativi e da precedenti giurisprudenziali - allo scopo di costruire modelli statistici finalizzati al raggiungimento di plurimi obiettivi, utili per un giurista, tra i quali quello, per l’appunto, di prevedere anche il possibile esito di un giudizio”.

⁴⁵ Nonostante la peculiarità sostanziale di questo ordinamento sovranazionale, in cui l’assenza di un vero e proprio “potere legislativo intestato ad un soggetto determinato, ma più ancora l’estrema limitatezza del diritto, che comunque nasce alla stregua di leggi da interpretare, ha tramutato la Corte (di Giustizia), quando essa formalmente interpreta la norma positiva europea, in quello che potremmo chiamare, con il linguaggio iperrealistico dei funzionari di Bruxelles, il quarto pilastro – insieme alla Commissione, al Parlamento e al Consiglio – della costruzione normativa dell’Unione”, C. Castronovo, *L’eclissi del diritto civile*, Giuffrè, 2015, p. 228. Pone l’accento sulla tendenza diffusa del giudicare, con “*infiniti richiami di che occupano intere pagine ed hanno il sopravvento sul diritto positivo, come fossimo in un sistema di common law*”, alla quale la “*giustizia digitale*” potrebbe porre un limite, riorientando il giudizio in chiave nuovamente ricognitiva piuttosto che volitiva, H. Simonetti, *Brevi note sul possibile uso dell’intelligenza artificiale nel processo amministrativo*, in Giustizia-amministrativa.it.

⁴⁶ Da ultimo, Cons. Stato, Sez. II, 13 gennaio 2025, n. 217 “l’art. 12 delle disposizioni preliminari al codice civile, quale “norma sulle norme”, sancisce in modo univoco il primato dell’interpretazione testuale sugli altri criteri ermeneutici, coerentemente con l’oggettiva evidenza che il quadro ordinamentale è costituito da proposizioni formali, i cui enunciati sono espressi in formulazioni linguistiche e il cui fine è rendere intellegibile la portata descrittiva e precettiva delle regole poste, tramite l’univocità della relazione tra il loro significante e il loro significato, garantendo, per tal via, la certezza del diritto. Ne discende che la chiarezza del testo della disposizione e l’assenza, sul piano dell’interpretazione letterale, di dubbi sulla derivante norma da essa estratta neutralizzano gli altri canoni interpretativi, i quali, invero, possono avere rilievo soltanto ove l’esegesi testuale offra un risultato ambiguo” con commento di L. Viola, in Lanuovaproceduracivile.com, 23 gennaio 2025, cfr. anche Cons. Stato, Sez. VI, 2 ottobre 2023 n. 8610.

⁴⁷ Nel panorama emergenti delle start up nazionali in ambito Legal Tech, si muove in questa direzione, ad esempio, *GiuriMatrix*, una piattaforma di IA, per ora specializzata solo nel diritto e processo civile, progettata e sviluppata da giuristi italiani, basata su LLM coadiuvati da RAG, ma che integra nella sua architettura algoritmica anche l’art. 12 delle Preleggi. Cfr. *Indagine conoscitiva sull’intelligenza artificiale presso la X Commissione della Camera dei Deputati*, 1 febbraio 2024 “*GiuriMatrix è una intelligenza artificiale generativa, addestrata per il diritto, le cui risposte sono individuate nella legge con citazione della fonte. La dottrina e la giurisprudenza vengono utilizzate, eccezionalmente, come mezzi per trovare la risposta*”. Cfr. anche <https://www.agendadigitale.eu/documenti/giustizia-digitale/giurimatrix-unia-imparziale-per-il-settore-legale/>

6. Orizzonti applicativi dei LLM nella Giustizia Amministrativa

Abstract: I LLM offrono potenzialità significative per la giustizia amministrativa e opportunità di trasformazione di processi complessi come la ricerca delle fonti normative, l'analisi dei fascicoli digitali, e la redazione di provvedimenti giudiziari. Consentono ricerche personalizzate, sintesi mirate e il confronto tra documenti giuridici, migliorando efficienza e precisione. Nel supporto alla redazione di atti, i LLM possono identificare criticità stilistiche e favorire chiarezza e sinteticità, ampliando anche la comprensibilità delle decisioni per un pubblico non tecnico. Infine, nei metodi di risoluzione alternativa delle controversie (ADR), offrono strumenti per l'automazione delle analisi preliminari e la simulazione di scenari negoziali, segnando una possibile svolta nelle modalità di gestione stragiudiziale delle dispute.

Sui rischi e i timori legati all'IA Generativa e ai LLM nella giustizia c'è un ampio consenso, che alimenta anche una comprensibile “*resistenza culturale*” all'innovazione. Di fatto, le perplessità circa l'uso delle tecnologie di IA nel campo della giustizia ruotano intorno al rischio che la decisione sia delegata alla macchina e che l'uomo-giurista si trasformi inesorabilmente in un mero esecutore di “pacchetti” algoritmici.

Siamo forse l'ultima generazione ad aver vissuto la Rivoluzione digitale, perché, come è stato osservato, “*le generazioni future non sapranno mai com'era una realtà esclusivamente analogica, offline, predigitale*”⁴⁸: Questo rende ancora più complesso immaginare le potenzialità dell'IA Generativa⁴⁹, le cui opportunità dovrebbero emergere attraverso *sperimentazioni*, secondo l'approccio innovativo della *regulatory sandbox*⁵⁰, coinvolgendo il maggior numero possibile di attori.

6.1 La ricerca “intelligente”

Come dimostrato dall'emergente mercato del *Legal Tech* al quale si è fatto sopra cenno, uno degli ambiti di maggiore applicazione dei LLM è quello dell'attività di ricerca e interpretazione delle fonti normative, dei precedenti giudiziari, della dottrina. Sistemi di ricerca “intelligente” sarebbero in grado, se accuratamente progettati e sviluppati con il supporto di esperti di

⁴⁸ L. Floridi, *L'etica dell'Intelligenza Artificiale. Sviluppi, opportunità, sfide*, Raffaello Cortina editore, 2022, p.11

⁴⁹ Per i progetti già in corso nella Giustizia amministrativa, che non rientrano nei casi d'uso considerati ad alto rischio secondo l'AI Act, in quanto volti a “*fornire esclusivamente supporto nello svolgimento di attività preparatorie, organizzative o di approfondimento*”, da ultimo la Giustizia Amministrativa è stata insignita dell'autorevole premio Agenda Digitale 2024. I casi d'uso e lo spirito che li muove sono descritti in B. Bruno, N. Bardini, F. Sivilli, *Intelligenza artificiale e Giustizia amministrativa: strategie di impiego, metodologie e sicurezza*, in www.giustizia-amministrativa.it, 2024. I progetti si inseriscono nella misura PNRR 1.6.5 (“*Digitalizzazione delle grandi PAC – Consiglio di Stato*”) e, come analiticamente riportato nel report, poggiano su un terreno fertile preparato da tempo e costituito da un livello di digitalizzazione molto elevato grazie al PAT dal 1 gennaio 2017; da scelte infrastrutturali adeguate, ispirate alla direzione strategica nazionale del *cloud first*; e, elemento cruciale, dalla “*forte integrazione delle competenze*”.

⁵⁰ La letteratura sui cd. *spazi di sperimentazione normativa* non è annoverabile in questa sede, essendo già sterminata. Per una ricognizione generale delle questioni e ampi riferimenti bibliografici, cfr. P. Longo, G. Centemero, *Regulatory Sandbox: analisi e prospettive parlamentari*, 2024, Maggioli Editore. Sia consentito il rinvio a G. Lo Sapia, *Il regolatore alle prese con le tecnologie emergenti. La regulatory sandbox tra principi dell'attività amministrativa e rischio di illusione normativa*, in Federalismi.it, 16 novembre 2022. Sulla rivisitazione del “metodo scientifico” di ispirazione galileiana anche per lo studio delle opportunità offerte dall'innovazione tecnologica dell'IA Generativa, cfr. le autorevoli osservazioni della dottrina amministrativista: “*Il diritto, per questo suo dispiegarsi tra teoria e applicazione nel concreto verificata dall'esame del comportamento dei consociati e dalla giurisprudenza, potrebbe essere assimilato al metodo galileiano. I principi, pur mirabilmente elaborati dalla dottrina, se non sono osservati dai soggetti dello ordinamento e se la giurisprudenza non li segue e non li applica nel concreto, rendendoli vincolanti, sono teorie (ipotesi) non verificate nell'esperienza, con conseguente loro abbandono perché non in grado di disciplinare l'esperienza giuridica. La società e la giurisprudenza validano la tesi prospettata dalla dottrina se la fanno propria; se ciò non avviene, la tesi non è utile, se non addirittura infondata*”, E. Follieri, *Il metodo nel diritto amministrativo in alcuni ordinamenti stranieri*, *Relazione introduttiva*, in www.giustizia-amministrativa.it, 2024.

dominio, di supportare il magistrato nella complessa individuazione delle norme applicabili al caso concreto, individuando eventuali antinomie tra disposizioni di rango e fonte diversa ed offrendo all'interprete la valutazione critica delle diverse opzioni risolutive proposte.

Ai LLM non sfuggirebbero le evoluzioni, anche semantiche, prodotte dalla bulimia legislativa, rendendo la ricerca più rapida, precisa e mirata. Tali modelli, debitamente addestrati, potrebbero semplificare la ricerca normativa combinando diverse funzionalità: indicizzazione intelligente delle fonti giuridiche in base a criteri indicati da chi formula la domanda in linguaggio naturale, correlando le norme con i precedenti giurisprudenziali e la dottrina; prospettazione di opzioni interpretative, secondo i criteri ermeneutici positivizzati, e sulla base della specifica questione giuridica da approfondire; comparazione semantica tra diverse formulazioni normative, anche sovranazionali. La differenza tra la ricerca tradizionale, fatta sui "data base" e quella supportata dai LLM può essere spiegata sotto diversi profili, ma non può essere davvero compresa, fino a quando tali modelli non si utilizzino nella pratica. I LLM offrono infatti un'interazione personalizzata e adattiva, sfruttando dati storici e memoria delle precedenti conversazioni con il singolo interlocutore, nel rapporto *one-to-one*; possono interpretare il significato del testo, riconoscere sinonimi, contesti e riferimenti normativi impliciti, riducendo il rischio di ottenere risultati irrilevanti o di perdere documenti rilevanti, solo perché non è stata inserita nella domanda la "giusta" parola chiave.

Ad esempio ad un LLM può essere formulato questo *prompt* in linguaggio naturale: "quali sono gli orientamenti più recenti della giurisprudenza amministrativa in materia di risarcimento del danno?"; invece, per soddisfare lo stesso interesse conoscitivo, con un sistema di ricerca tradizionale, deve essere indicata la parola chiave: "responsabilità civile della PA", "2043 c.c." "30 c.p.a.". Soprattutto, i LLM offrono, oltre che l'elenco delle fonti individuate, direttamente una sintesi dei loro contenuti, evidenziando punti chiave e divergenze tra i diversi risultati, suggerendo incongruenze, che alimentano a loro volta il pensiero critico e la valutazione con intelligenza umana del materiale. I sistemi tradizionali invece, come noto, forniscono un elenco di documenti che l'utente deve leggere e sintetizzare autonomamente.

Restando nelle metafore *umanizzate*, potrebbe dirsi che il data-base tradizionale, digitalizzato sulla base di criteri non noti all'utilizzatore, è un po' come una biblioteca con un catalogo di libri predefinito, di cui però è necessario conoscere *ex ante* almeno il titolo ("parola chiave"); un sistema IA, invece, è come un bibliotecario esperto che prima di tutto sa capire cosa si sta cercando e poi sa individuare e suggerire il libro giusto e letture correlate, indicando anche i capitoli più rilevanti da leggere con più attenzione.

6.2. Analisi dei documenti e degli atti del fascicolo digitale

E' indubbio che nel processo amministrativo, nonostante il dichiarato principio della atipicità della prova, l'istruttoria sia fondata di fatto su prove documentali "precostituite", formate, per lo più, nel procedimento amministrativo (con documenti oramai quasi sempre nativi digitali) e redatti, per la maggior parte, in linguaggio naturale, pur non escludendosi rappresentazioni documentali per immagini (planimetrie, progetti allegati a richieste di permesso di costruire, fotografie satellitari per la determinazione della data di realizzazione di un abuso edilizio, mappe catastali, grafici allegati ai piano regolatori; immagini generate da applicazioni digitali

come Google Earth, Google Maps, Google Street View, già in uso nei giudizi amministrativi⁵¹); una mole di documenti – e immagini - versata nel singolo fascicolo processuale, anche grazie anche alla facilità del deposito consentito dalla digitalizzazione, tanto che sono diffuse anche le inutili duplicazioni.

Se adeguatamente addestrato, un LLM avrebbe l'abilità, in pochi secondi di tempo, di individuare le parti dei documenti che contengano riferimenti all'informazione specifica richiesta con il *prompt* (es. *“rispetto alla seconda censura del ricorso per motivi aggiunti, quali sono gli atti dell'amministrazione e del controinteressato che contengono le corrispondenti eccezioni difensive?”*); di estrarre automaticamente dai documenti ed atti processuali, informazioni chiave come date, norme citate, precedenti giurisprudenziali; di generare sintesi delle diverse argomentazioni logico-giuridiche, rispetto ad una specifica questione; di riorganizzare, sulla base del contenuto dei documenti, l'andamento cronologico del procedimento come risulta dalla ricostruzione delle parti e dai documenti versati in giudizio.

In sostanza, un LLM integrato nei flussi di lavoro può accelerare e migliorare l'analisi documentale, muovendosi nel mare di informazioni riversate nel fascicolo con velocità inesigibili per la mente umana, mantenendo elevati standard di precisione e affidabilità, mediante il richiamo alla specifica parte del documento, suscettibile così di verifica da parte del giudice. Nella sperimentazione di IA ai fascicoli processuali, ha ad esempio avuto una certa eco mediatica, la notizia relativa alla Procura Generale di Perugia che ha impiegato l'IA per affrontare l'enorme mole documentale legata a un caso giudiziario complesso, quale quello della revisione della condanna per calunnia di un ex collaboratore di giustizia. L'archivio contava circa 140 mila pagine e un sistema di IA è stato in grado di estrapolare, in 40 secondi, solo i documenti contenenti le contestazioni, secondo la specifica richiesta formulata.

In sintesi, un sistema basato su un LLM correttamente progettato e addestrato, avrebbe l'abilità di cooperare nella ricerca documentale con il giudice, fungendo da “assistente documentale intelligente”, da integrare nell'Ufficio del Processo; non volto a sostituire l'attività critica del magistrato, ma a supportarla ampliando l'orizzonte dell'analisi informativa, garantendo maggiore efficienza nella ricerca, riduzione della ridondanza documentale e una più rapida comprensione dell'istruttoria. Non un ausiliario ex art. 21 c.p.a., appunto, da far intervenire in modo eventuale; ma un assistente digitale, disponibile h24, addestrato nello specifico settore giuridico di riferimento (o della specifica Sezione), capace di individuare, organizzare e richiamare puntualmente le informazioni essenziali. Un supporto non progettato ed addestrato per sostituire il ruolo del giudice, per potenziarne l'efficienza conoscitiva e le abilità analitiche, permettendogli di dedicare più tempo all'attività critica e decisionale, e, in ultima analisi, a soffermarsi sulle specificità dei casi concreti, che reclamano giustizia.

6.3. Chiarezza e sinteticità degli atti con il supporto di IA

È noto che questo rappresenta un ambito di applicazione particolarmente sensibile, dove il dibattito politico è acceso, alimentato dalla paura del cosiddetto *'giudice-robot'*. Ma potrebbe

⁵¹ Cfr L. Tarantino, *Istruttoria nel processo amministrativo alla prova delle nuove tecnologie*, in L. R. Perfetti e G. Pesce, *L'istruttoria nel processo amministrativo*, Giuffrè, 2024, p. 139 *“E' interessante notare come anche il processo amministrativo sia influenzato dall'evolversi delle nuove tecnologie e come ciò dipenda non solo dal fatto che anche nel nostro giudizio si rispecchia la realtà di una società sempre più conformata dalle tecnologie digitali, ma anche in forza del fatto che il processo amministrativo si avvale, ed è in predicato di avvalersi, di alcune di queste tecnologie”*.

essere utile guardare alle opportunità offerte dalle tecnologie dell'**IA Generativa** da una prospettiva diversa.

I sistemi di IA basati su LLM, infatti, potrebbero essere impiegati, nell'ambito delle eccezioni previste dall'art. 6, paragrafo 3, dell'AI Act, nella fase successiva alla redazione della 'bozza' del provvedimento giurisdizionale, da parte del giudice al fine di individuare punti critici o non facilmente comprensibili; migliorare il grado di chiarezza e comprensibilità dell'impianto complessivo della motivazione; evidenziare frasi involute o inutilmente complicate, in modo da agevolarne la lettura anche da parte delle parti sostanziali del giudizio⁵²; estrapolare i principi giuridici, anche non espressamente indicati nella motivazione; conformare la redazione della decisione al principio di chiarezza e sinteticità di cui all'art. 3 c.p.a. che è imposto, paritariamente, sia ai difensori che ai magistrati⁵³.

L'obiezione più immediata è che, una volta accettata l'integrazione di un LLM nel processo di redazione, diventa difficile tracciare un confine chiaro tra un intervento ex post e uno ex ante. Tuttavia, se non fosse una rivoluzione, non ci sarebbero sfide difficili da affrontare, per le quali nessuno ha soluzioni in tasca. Peraltro, l'idea di un divieto assoluto all'uso degli LLM nella redazione della sentenza *tout court*, sostenuta da molti, risulta non solo incompatibile con il diritto sovranazionale—sollevando quindi potenziali questioni di disapplicazione delle norme nazionali—ma anche in contrasto con la realtà di una diffusione ormai incontrollata di questi modelli che, se non adeguatamente e tempestivamente governata, genera un nuovo rischio emergente legato all'IA: la cosiddetta **Shadow AI** sulla quale ci si sofferma nelle conclusioni.

6.4. *La vis espansiva dell'IA nelle ADR*

L'applicabilità dei LLM ai metodi di risoluzione alternativa delle controversie (ADR) offre un terreno ricco di prospettive. Ad essi, accanto all'attività giurisdizionale, si riferisce anche il punto 8 dell'Allegato III dell'AI Act, sopra menzionato. Si tratta di un ambito in cui le tecnologie innovative di IA intercettano le recenti evoluzioni normative, che mirano a rinforzare strumenti alternativi per la soluzione delle controversie, anche in funzione deflattiva del contenzioso giudiziale. Basti considerare che, nella vasta area delle ADR, rientra anche l'istituto del Collegio consultivo tecnico di cui all'art. 215 del decreto legislativo 31 marzo 2023, n. 36⁵⁴. Nel mercato dei *tool* di IA Generativa applicati al settore legale, una fetta in rapida espansione sono i LLM progettati e sviluppati con abilità di "negoiazione", di

⁵² L'ausilio dei LLM al fine di attuare nella pratica i principi di chiarezza e sinteticità potrebbe peraltro espandersi anche fuori dalle aule giudiziarie, al fine di rendere accessibile, in un linguaggio meno tecnico e più comprensibile anche ai "non addetti ai lavori", il significato della decisione, specie per quelle ad elevato impatto sociale o economico, in modo da facilitare il lavoro degli "uffici stampa", sotto la guida del Collegio che però ha adottato la specifica decisione.

⁵³ Nella redazione degli atti del processo, da parte dei difensori, da ultimo la giurisprudenza amministrativa ha osservato come il superamento dei limiti dimensionali, corollari del fondamentale principio di sinteticità, costituisca "questione di rito afferente all'ordine pubblico processuale, stabilito in funzione dell'interesse pubblico all'ordinato, efficiente e celere svolgimento dei giudizi, ed è rilevabile d'ufficio a prescindere da eccezioni di parte" Cons. Stato, Sez. II, 17 settembre 2024, n. 7614.

⁵⁴ Con una funzione consultiva anche più ampia, di quella meramente "risolutiva", da esercitarsi al fine di dirimere, anche in via preventiva, questioni concrete che potrebbero ostacolare la celere esecuzione del contratto, in ossequio al principio chiave di risultato; il CCT si ispira infatti al più noto istituto del "Dispute Review Board" di derivazione anglosassone, finalizzato a dirimere in modo celere ed efficace i problemi, che possono concernere sin dall'inizio della esecuzione il rapporto contrattuale e, quindi, condizionare o bloccare i lavori. Si consideri altresì che la figura del Dispute Board è ben conosciuta nell'esperienza della contrattualistica internazionale e mediante esso si intende individuare una soluzione più duttile e flessibile per risolvere il problema del complessivo ingolfamento della macchina processuale e per migliorare l'efficienza degli appalti pubblici", M. Macchia, *Alla ricerca di esperti... Rimedi alternativi ed esecuzione dell'appalto*, in *Giornale di diritto amministrativo*, n. 1, 2024, p. 19

automazione dell'analisi preliminari delle diverse posizioni delle parti, di simulazione degli scenari per formulare proposte negoziali finalizzate a soluzioni condivise dalle parti⁵⁵. Da ultimo, la decisione del Consiglio dell'UE, assunta in data 25 settembre 2024, sulla nuova Direttiva UE sulle ADR, di dismettere la piattaforma ODR (*Online Dispute Resolution*), che era destinata a risolvere stragiudizialmente e in modalità virtuale controversie tra consumatori e pubbliche amministrazioni, chiude un capitolo fallimentare, ma apre anche nuovi orizzonti all'applicazione di tecnologie innovative in un settore emergente. Tanto che proprio all'AI fanno espresso riferimento anche le recenti Linee Guida sulle ODR del CEPEJ⁵⁶ e che cominciano a proliferare ricerche specifiche sul tema⁵⁷. Nella giustizia amministrativa, i metodi di risoluzione stragiudiziale delle controversie non oggi hanno molto spazio applicativo, ma la disponibilità di ausili tecnologici avanzati, debitamente progettati e poi accuratamente applicati, potrebbe rappresentare l'abbrivio per valorizzare con approcci innovativi anche questo ambito, annoverato da autorevole dottrina tra le sfide del processo amministrativo nel futuro prossimo⁵⁸.

7. Conclusioni: il rischio inedito della Shadow AI

Abstract: L'IA generativa (Gen AI) e i LLM rappresentano una rivoluzione per il diritto e la giustizia, offrendo strumenti capaci di supportare il lavoro giuridico con una nuova "intelligenza collettiva". Tuttavia, sollevano interrogativi sulla progettazione responsabile, il prompt engineering, e la formazione per bilanciare fiducia e supervisione, evitando l'automation bias. All'orizzonte emerge il rischio della Shadow AI, ossia l'uso non trasparente di LLM non specificamente addestrati per il diritto amministrativo, con implicazioni su sicurezza, privacy e bias decisionali. Serve un approccio regolatorio bilanciato, capace di mitigare i rischi senza soffocare le opportunità offerte da questa tecnologia rivoluzionaria.

L'IA Generativa, attraverso i LLM, costituisce l'emblema della Rivoluzione digitale oggi e non può che impattare sia sulle professioni giuridiche che sul complessivo sistema della giustizia, trasformando competenze, processi e dinamiche di interazione tra esseri umani e ausili digitali. Tuttavia, l'introduzione di tali tecnologie solleva interrogativi fondamentali e specifici rispetto al dominio legale: come conformare, fin dalla fase della progettazione, i LLM, affinché diventino "copiloti" efficaci in ordinamenti giuridici di *Civil Law*; quali competenze multidisciplinari sono necessarie perché siano addestrati e adattati alla giustizia amministrativa, in modo responsabile; come interagire in modo consapevole con tali modelli, in modo da assicurare che il rapporto dialogico non si traduca in una acritica validazione delle risposte fornite dall'IA; in definitiva, come assicurare, anche sul piano concreto dell'attività giurisdizionale, il delicato equilibrio tra la fiducia e la supervisione umana, evitando che la prima si tramuti in cieco affidamento e la seconda in un controllo eccessivo, neutralizzando di fatto gli eventuali benefici derivanti dalle innovazioni tecnologiche.

In gioco c'è una nuova visione dello *Human-In-The-Loop*, in cui i LLM, se ben progettati e guidati, potrebbero trasformarsi in assistenti riflessivi e perspicaci, capaci di affiancare il giudice e trasformare l'organizzazione dell'*Ufficio del Processo*⁵⁹, attraverso una nuova

⁵⁵ <https://www.smartsettle.com/>. Per una aggiornata lettura del fenomeno, cfr. arXiv:2410.07053

⁵⁶ *Guidelines on Online Alternative Dispute Resolution*, CEPEJ(2023)19Final, adottate il 3 dicembre 2023.

⁵⁷ Cfr. F. Amato, M. Fonisto, M. Giacalone, C. Sansone, *An Intelligent Conversational Agent for the Legal Domain*, in *Information*, 2023, 6, 307.

⁵⁸ M. Clarich, *Pubblica amministrazione e trasformazioni della giustizia amministrativa*, in *Giornale di diritto amministrativo*, n. 6, 2024, p. 781

⁵⁹ Come osservato, si tratta di un progresso graduale, del quale non si possono prevedere con certezza gli esiti, neanche in relazione all'attività amministrativa, "non è escluso che in futuro una decisa propensione per la informatizzazione amministrativa possa portare a talune modifiche sulle strutture procedurali: questo è auspicabile che accada soprattutto quando si implementi l'uso di strumenti di informatica decisionale più potenti e «autonomi» dall'uomo rispetto a quelli sino

dimensione di “intelligenza collettiva”, coordinata dal giudice, adeguatamente formato e pienamente consapevole dei limiti e dei rischi di questa dirompente tecnologia.

Tuttavia, mentre il tempo necessario per sviluppare policy adeguate scorre inesorabile, l’urgenza di comprendere i limiti e le opportunità dell’IA cresce esponenzialmente. L’anno 2025 si apre con previsioni, che nessuno può effettivamente stabilire quanto affidabili, sui trend trasformativi: l’ascesa degli **Agenti IA**, l’esplosione di modelli “*small*” ed *open source*, il focus sulle abilità di *reasoning step-by-step*, la sinergia tra AI e computer quantistici.

Questi sviluppi promettono maggiore efficienza e adattabilità, ma, per la velocità con cui si presentano, portano con sé anche rischi inediti, tra cui il dilagare della cd. *Shadow AI*⁶⁰, ossia il rischio di un uso pervasivo, non regolato e tecnicamente impreparato degli LLM da parte dei singoli, nell’esercizio dell’attività lavorativa, senza la supervisione delle strutture organizzative cui spetta di indicare le policy più adeguate. L’uso solitario di sistemi dilaganti, di facile uso, magari inizialmente sperimentati in ambiti privati e poi riversati nello svolgimento dell’attività lavorativa, potrebbe sì far acquisire consapevolezza e dimestichezza con i nuovi “strumenti”, ma anche minare i principi fondamentali del giusto processo, con ricadute su sicurezza, privacy e imparzialità decisionale. Affrontare questa realtà richiede regole chiare, calibrate sulle specificità del contesto giuridico, ma anche tempestive e proattive, non solo volte a chiudere porte, ma anche ad aprirne. Non possiamo permetterci né inerzia, né un approccio meramente difensivo: il futuro dell’IA nel diritto non è più un’ipotesi, ma una traiettoria già tracciata, su cui la tecnologia avanza a ritmo incalzante.

*Glossario**

Agenti IA - Un agente IA è un sistema di IA in grado di percepire l’ambiente, prendere decisioni e agire in modo autonomo per raggiungere determinati obiettivi. Gli agenti possono variare da semplici script di automazione a modelli avanzati come gli autonomous AI agents che apprendono e si adattano, eseguendo operazioni sulla base di modelli di apprendimento. Anche chatbot avanzati possono essere considerati Agenti IA, se comprendono il contesto e sono abilitati ad agire di conseguenza.

ad oggi già adoperati dalle amministrazioni pubbliche italiane. E però, ancorché l’informatizzazione rechi con sé inevitabili mutamenti appare opportuno che tali cambiamenti – perché possano essere più facilmente accettati e meglio valorizzati da parte dei cittadini, oltre che da parte delle stesse amministrazioni – comportino un adattamento dei vecchi istituti alle nuove situazioni, piuttosto che uno stravolgimento delle classiche modalità di organizzazione dell’attività e del funzionamento delle p.a.” A.G. Orofino, La semplificazione digitale, in Il diritto dell’economia, 3, 2019, pp. 87-112.

⁶⁰ <https://www.rivista.ai/2024/10/20/shadow-ai-cose-e-come-gestire-lai-ombra-lintelligenza-artificiale-non-ufficial>

*Il **Glossario** contiene descrizioni delle locuzioni tecniche utilizzate nel contesto trattato, con l’obiettivo di renderne comprensibile il significato, con riferimenti anche normativi, dove esistenti. Tuttavia, non assume il carattere di un documento tecnico in senso stretto. La sua redazione è stata supportata dal confronto con esperti del settore, tra cui l’Ing. Mario Ettore, Chief Information Security Officer e Fraud Management Officer presso l’Agenzia nazionale per l’attrazione degli investimenti e lo sviluppo di impresa (InvItalia s.p.a.)

Black Box - Nell'ambito dell'IA, il termine **Black Box** (scatola nera) si riferisce a modelli o algoritmi il cui funzionamento interno non è direttamente interpretabile o comprensibile dall'utente. In altre parole, pur essendo in grado di produrre risultati accurati, questi sistemi non permettono di spiegare in modo trasparente come e perché una determinata decisione è stata presa. Le reti neurali profonde (Deep Learning) e i modelli di AI avanzati, come i Large Language Models (LLM), sono spesso considerati Black Box perché operano attraverso milioni o miliardi di parametri che interagiscono in modi complessi e non immediatamente interpretabili. La Black Box è citata nel Libro Bianco sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia (COM(2020) 65 Final del 19 febbraio 2020, come sinonimo di "opacità". Non si ritrova invece nell'AI Act, sostituito da questo sinonimo (Cons. 61)

ChatGPT - ChatGPT è un chatbot, disponibile nella sua ultima versione a pagamento e con interfaccia web, basato su un LLM denominato GPT (Generative Pre-trained Transformer GPT). Il modello GPT è stato sviluppato da OpenAI e progettato per generare testi in linguaggio naturale in modo coerente e contestualizzato, quale sistema avanzato di elaborazione del linguaggio naturale (NLP), addestrato su vasti insiemi di dati testuali per comprendere e rispondere a domande, generare contenuti, riassumere informazioni e interagire con gli utenti in modo conversazionale. Il primo modello è GPT-1 e risale al 2018, con 117 milioni di parametri. Nel 2020 è stato sviluppato il modello GPT-3, con 175 miliardi di parametri e la sua versione ottimizzata (GPT 3.5) è stata alla base di ChatGPT, lanciato sul mercato di massa il 30 novembre 2022. L'attuale modello in uso è GPT 4 multimodale, capace di analizzare immagini oltre al testo. La società proprietaria non ha fornito informazioni sul numero di parametri utilizzati dell'ultimo modello. ChatGPT oggi raccoglie circa 250 milioni di utenti settimanali e il suo modello di base GPT è il motore di altre numerose applicazioni, anche note, tra cui DuolingoMax, con funzioni di tutor per imparare una nuova lingua; Be My Eyes, progettata per rendere il mondo virtuale più accessibile alle persone ipovedenti; Khan Academy, piattaforma di corsi di formazione online gratuiti; Copilot di Microsoft, la cui icona compare in alto a destra su tutti gli schermi dei PC in dotazione ai magistrati amministrativi, anche quando si accede a Word..

Deep Learning - Il Deep Learning è un metodo di apprendimento automatico (Machine Learning) che utilizza **reti neurali artificiali** con più livelli (deep neural networks) per elaborare e riconoscere schemi complessi nei dati. Il Deep Learning non richiede pertanto che un programmatore definisca esplicitamente le regole per analizzare i dati: il modello basato su questa struttura algoritmica impara autonomamente attraverso l'addestramento su grandi quantità di informazioni, che fungono da "esempi". Geoffrey Hinton è considerato il "padre del Deep Learning" ed è stato insignito, insieme a John Hopfield, del Premio Nobel per la fisica 2024 dall'Accademia reale svedese delle scienze "per scoperte e invenzioni fondazionali che abilitano l'apprendimento automatico nelle reti neurali artificiali". La locuzione Deep Learning fu coniata dallo Hinton, in una conferenza a Vancouver nel 2007, il giorno del suo sessantesimo compleanno.

General Purpose Model - I modelli di intelligenza artificiale general-purpose (General-Purpose AI Models) sono modelli avanzati di machine learning progettati per svolgere un'ampia gamma di compiti, senza essere specializzati in una sola funzione o in un solo dominio. A differenza dei modelli tradizionali, che vengono addestrati per uno specifico compito (es. riconoscimento facciale, analisi delle immagini mediche), i modelli general-purpose possono essere adattati a molteplici applicazioni con minimo o nessun fine-tuning. In questa categoria rientrano sia i LLM che i LMM. La loro definizione e disciplina non era prevista nella proposta originaria della Commissione europea dell'AI Act del 2021. Nel testo vigente un modello di IA per finalità generali è definito nell'art. 3 n. 63 come "un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, ad eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato" e tali modelli sono disciplinati dal Capo V che distingue tra quelli a rischio sistemico e quelli ritenuti privi.

GPU (Graphics Processing Unit): GPU è un processore specializzato progettato per eseguire in parallelo un numero elevato di operazioni matematiche, particolarmente quelle legate all'elaborazione grafica e al calcolo vettoriale. Le GPU sono fondamentali nell'intelligenza artificiale, perché la loro architettura è ottimizzata per eseguire in parallelo milioni di calcoli, un requisito essenziale per l'addestramento e l'inferenza delle reti neurali.

Human In The Loop - il concetto di Human-in-the-Loop (HITL) nell'intelligenza artificiale si riferisce a sistemi in cui l'intervento umano rimane essenziale in uno o più momenti del processo decisionale dell'IA. Questo approccio garantisce che l'automazione non operi in modo completamente autonomo, ma sia costantemente supervisionata, corretta e affinata dagli esseri umani.

Intelligenza Artificiale (IA)- L'Intelligenza Artificiale è un ramo delle scienze computazionali che progetta e sviluppa sistemi in grado di simulare processi cognitivi umani, come l'apprendimento, la percezione e la risoluzione di problemi. L'IA oggi utilizza tecniche avanzate di machine learning per analizzare dati, riconoscere schemi e prendere decisioni in modo autonomo o assistito. La definizione contenuta nell'AI Act si riferisce a “sistemi di IA” come ad “un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali” (art. 3 n. 1).

Intelligenza Artificiale Generativa (IA Gen) - L'AI Generativa è branca delle tecnologie di intelligenza artificiale progettata per creare nuovi contenuti, simili a quelli prodotti dagli esseri umani. Funziona imparando da enormi quantità di dati e utilizzando modelli avanzati per generare testi, immagini, musica, video e persino codice. A differenza di sistemi di IA tradizionali, che si limitano ad analizzare dati o prendere decisioni, l'AI Generativa può produrre informazioni “originali”, completare testi, creare immagini realistiche o simulare scenari. I suoi usi spaziano dalla scrittura automatica alla grafica digitale, dalla ricerca scientifica al supporto creativo. (rif. Considerando 105 Ai Act)

Large Language Model (LLM)- Un LLM è un modello di intelligenza artificiale basato su reti neurali profonde, progettato per comprendere, generare e manipolare il linguaggio naturale con un alto grado di coerenza e complessità. Questi modelli sono addestrati su enormi quantità di dati testuali e utilizzano l'architettura Transformer. Tra i più noti, GPT di Open AI, Bert di Google, LLaMA di Meta, PaLM di Google DeepMind, DeepSeek R1, della start up cinese DeepSeek, con 245 miliardi di parametri. Nell'AI Act, un riferimento ai “modelli linguistici di grandi dimensioni) quali esempi di modelli di IA per finalità generali è contenuto nell'allegato XIII, lett. d).

Large Multimodal Language (LMM) - Un Large Multimodal Language Model (LMM) è un modello avanzato di intelligenza artificiale progettato per elaborare e generare informazioni in più modalità di input e output, come testo, immagini, audio e video. A differenza dei tradizionali Large Language Models (LLM), che operano esclusivamente sul linguaggio testuale, gli LMM integrano diverse tipologie di dati, permettendo loro di comprendere immagini e generare descrizioni, interpretare audio, analizzare video e rispondere in modo più ricco e contestualizzato. Un esempio noto di LMM è GPT-4V (capace di interpretare immagini e testo) o Gemini di Google.

Legal Prompting - Il Legal Prompting è una disciplina emergente che si focalizza sull'arte di formulare istruzioni precise per ottenere risposte coerenti e pertinenti da modelli di intelligenza artificiale nel contesto legale, contestualizzando in questo dominio le tecniche di **Prompt Engineering**.

Legal Tech – il Legal Tech (Legal Technology) è l'uso di tecnologie avanzate per ottimizzare, automatizzare e migliorare i servizi legali e il funzionamento degli studi legali, delle aziende e delle istituzioni giudiziarie. Comprende software, piattaforme e anche applicazioni basate su intelligenza artificiale per rendere più efficienti le attività giuridiche, relative alla redazione automatizzata di documenti, alla ricerca giuridica, all'analisi dei contratti, alla risoluzione online delle dispute (ODR). Applicazioni diffuse soprattutto negli ordinamenti anglosassoni sono Ross Intelligence, DoNotPay, Kira

System. Nell'area in forte espansione degli "assistenti legali di IA", uno dei più diffusi oggi negli Stati Uniti è CoCounsel, definito come il primo "GenAI Legal assistant" fondato sul modello GPT 4.0 di OpenAI. Nel mercato italiano, tra le applicazioni emergenti, si annoverano GiuriMatrix e Lexroom.

Machine Learning -il Machine Learning (ML) è una branca dell'intelligenza artificiale che permette ai computer di apprendere dai dati ed effettuare previsioni o decisioni senza essere programmati esplicitamente. A differenza della programmazione tradizionale, dove ogni istruzione è scritta manualmente, nel ML il sistema riconosce schemi nei dati e migliora le sue prestazioni con l'esperienza. Quando si dice che un modello di Machine Learning "apprende", non significa che acquisisce conoscenza come farebbe un essere umano, ma che identifica pattern nei dati e li usa per fare previsioni o prendere decisioni in modo autonomo. L'apprendimento avviene attraverso l'ottimizzazione di una funzione matematica che permette al modello di migliorare le sue prestazioni nel tempo, riducendo errori e affinando le risposte.

Natural Language Processing (NLP) - Il NLP (analisi del linguaggio naturale) è un campo dell'intelligenza artificiale e della linguistica computazionale che si occupa di permettere ai computer di comprendere, interpretare, generare e rispondere al linguaggio umano in modo significativo. L'NLP combina modelli matematici, statistici e tecniche di *deep learning* per elaborare il testo o il parlato, estraendone informazioni, riconoscendo schemi e traducendo significati. Le principali applicazioni dell'NLP includono traduzione automatica, chatbot e assistenti virtuali, riconoscimento vocale, sintesi del linguaggio e classificazione dei documenti.

Prompt Engineering - Il Prompt Engineering è la pratica di progettare e ottimizzare input testuali per ottenere risposte precise ed efficaci dai modelli di intelligenza artificiale generativa, come i Large Language Models. I modelli di IA generativa rispondono in base al prompt, ovvero il testo di input fornito dall'utente. Scrivere un prompt efficace significa guidare il modello per ottenere il miglior risultato possibile. Alcune tecniche, come quella "Chain of Thought" prevedono la stimolazione del modello guidando passo dopo passo la risposta e sono già parzialmente integrate nei modelli più avanzati come GPT- 4 e PaLM 2.

Reti Neurali Artificiali - Le Reti Neurali Artificiali (Artificial Neural Networks - ANN) sono modelli matematici ispirati alla struttura e al funzionamento del cervello umano. Sono utilizzate nell'Intelligenza Artificiale per riconoscere schemi, apprendere dai dati e risolvere problemi complessi come la visione artificiale, la comprensione del linguaggio naturale e la generazione di contenuti.

Retrieval-Augmented Generation (RAG)- Il Retrieval-Augmented Generation (RAG) è una tecnica avanzata nell'intelligenza artificiale che combina il recupero di informazioni da fonti esterne con la generazione di testo. Questo approccio migliora l'accuratezza e l'affidabilità delle risposte fornite dai LLM, poiché consente di recuperare informazioni (Retrieval) pertinenti da una base di dati già individuata come affidabile e i qualità, integrando le conoscenze acquisite nella fase di addestramento.

Social Web - Il Social Web si riferisce all'evoluzione di Internet da una rete statica di contenuti pubblicati da pochi (Web 1.0) a un ecosistema interattivo e partecipativo (Web 2.0), in cui gli utenti creano, condividono e modificano contenuti in tempo reale. Il Social Web comprende piattaforme come social network (Facebook, Twitter/X, Instagram, LinkedIn), blog, forum, wiki e altre tecnologie collaborative, dove il valore principale è generato dalle interazioni tra gli utenti. Questa trasformazione ha reso il Web un spazio dinamico, basato sulla condivisione di informazioni, sulle reti sociali digitali e sulla partecipazione collettiva, influenzando profondamente la comunicazione, il giornalismo, l'economia e persino la politica.

Transformer - Il Transformer è un tipo di modello di intelligenza artificiale progettato per elaborare e comprendere grandi quantità di dati sequenziali, come testi, immagini o suoni, in modo rapido ed efficiente. Introdotto nel 2017 da un gruppo di ricercatori di Google ("Attention is All You Need"), ha rivoluzionato il modo in cui i sistemi di AI gestiscono il linguaggio naturale e altre forme di informazione

strutturata. La sua caratteristica principale è il meccanismo di attenzione (self-attention), che permette al modello di analizzare il significato di ogni parola o elemento di una sequenza considerando il contesto complessivo, invece di elaborarlo parola per parola in ordine. Questo lo rende molto più veloce e preciso rispetto ai modelli precedenti. I Transformer sono alla base di modelli avanzati di AI, come GPT (usato in ChatGPT), BERT (utilizzato per il motore di ricerca di Google), la cui "T" sta ad indicare proprio tale modello.