

Tokenizzazione, Embedding ed Etica dell'IA

Quello che serve all'avvocato per
governare l'innovazione senza rischi.

Non entriamo nella matematica dell'IA, ma in ciò che serve all'avvocato per non sbagliare.
Un percorso dal funzionamento tecnico alla responsabilità deontologica.

Perché guardare
“sotto il cofano”?

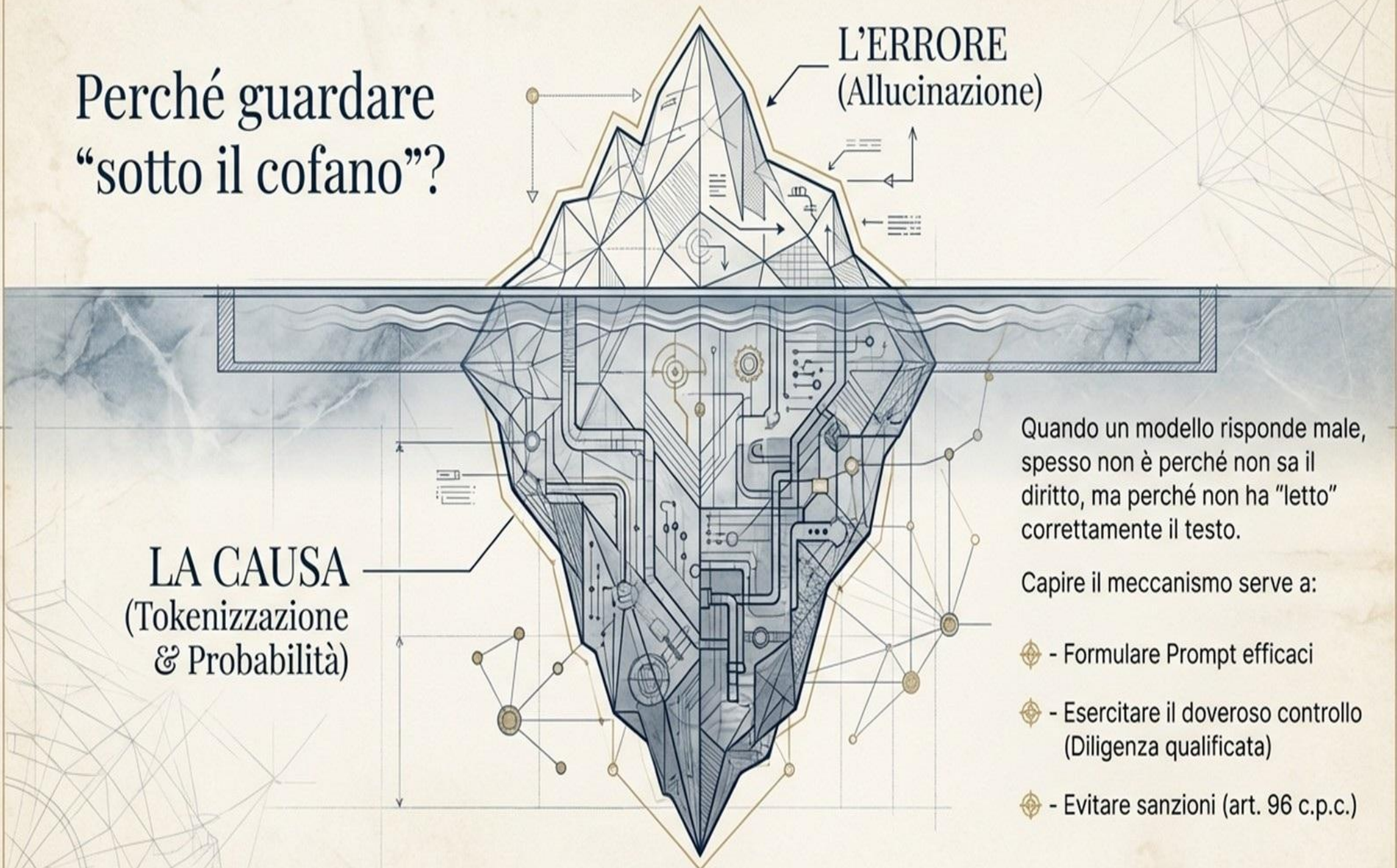
L'ERRORE
(Allucinazione)

LA CAUSA
(Tokenizzazione
& Probabilità)

Quando un modello risponde male,
spesso non è perché non sa il
diritto, ma perché non ha “letto”
correttamente il testo.

Capire il meccanismo serve a:

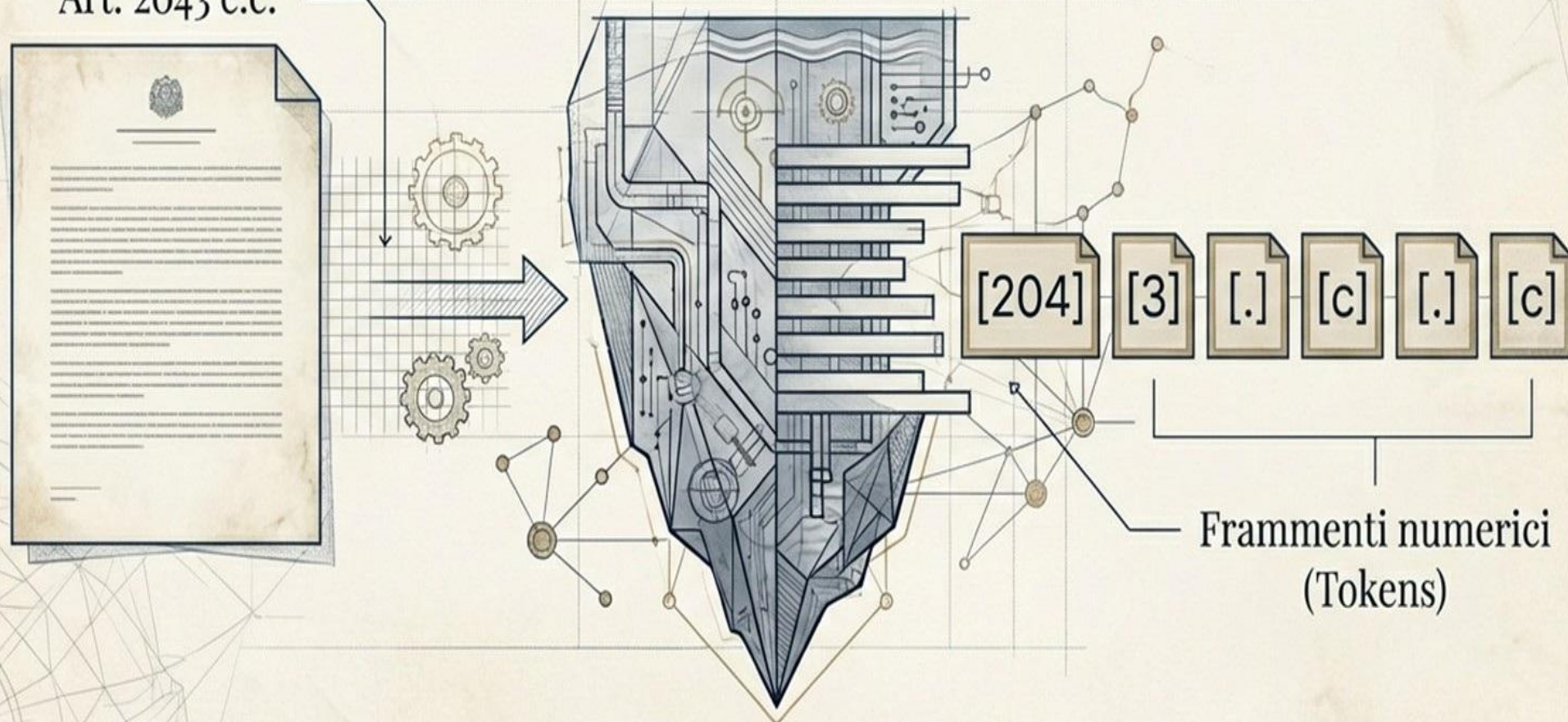
- ◆ - Formulare Prompt efficaci
- ◆ - Esercitare il doveroso controllo
(Diligenza qualificata)
- ◆ - Evitare sanzioni (art. 96 c.p.c.)



Cos'è la Tokenizzazione: Il “fascicolo smontato”

Il modello non “legge” parole, ma calcola frammenti numerici. Immaginate di smontare un fascicolo in singoli foglietti volanti numerati: il legame semantico si rompe in numeri.

Art. 2043 c.c.



Limiti di contesto in token delle principali soluzioni commerciali di IA

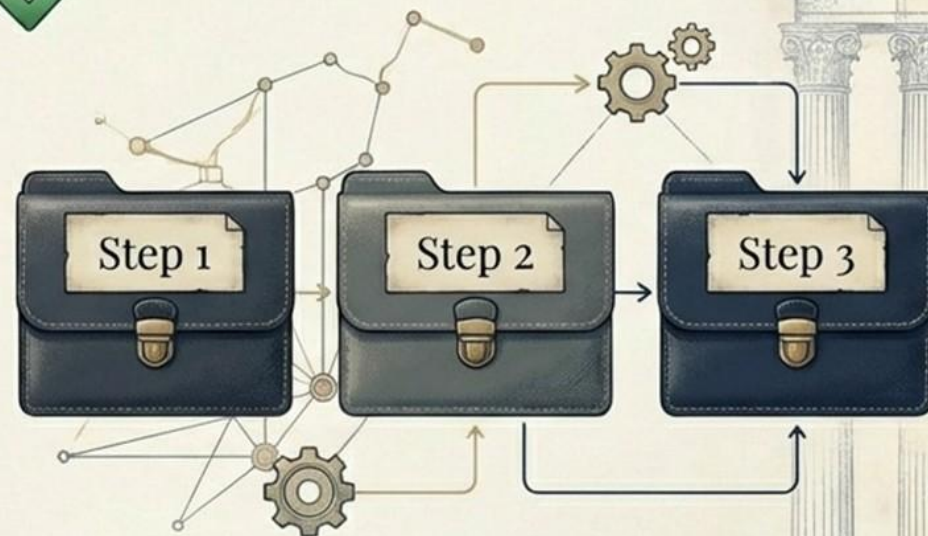
Provider	Modello principale	Contesto massimo teorico	Contesto tipico versioni commerciali	Soluzioni commerciali
OpenAI	GPT-5.x / GPT-4.1+ / GPT-4o	~400K token	Inferiore nelle app (Plus/Team)	ChatGPT Free, Plus, Team, Enterprise
Anthropic	Claude Opus 4.x / Sonnet 4.x	Fino a 1M token	~200K token	Claude Free, Pro, Team, Enterprise
Google	Gemini 3 Pro / Gemini 1.5 Pro	Fino a 1M token	Ridotto nelle interfacce web	Gemini Free, Advanced, Workspace AI
Meta	Llama 3.1	~128K token	Variabile	API cloud / open source
Mistral AI	Mistral Large / Mixtral	32K-128K token	Variabile	Le Chat, Enterprise API
Alibaba	Qwen series	128K+ token	Variabile	Tongyi Qwen / Cloud AI
Perplexity	Claude/GPT (aggregatore)	Dipende dal modello	Ridotto rispetto al teorico	Perplexity Pro

Implicazioni operative: Il "Drafting" per la macchina

Meglio tre richieste brevi che una richiesta enorme. Se l'avvocato è chiaro e divide il problema, l'IA sbaglia meno. La chiarezza espositiva diventa un requisito tecnico.

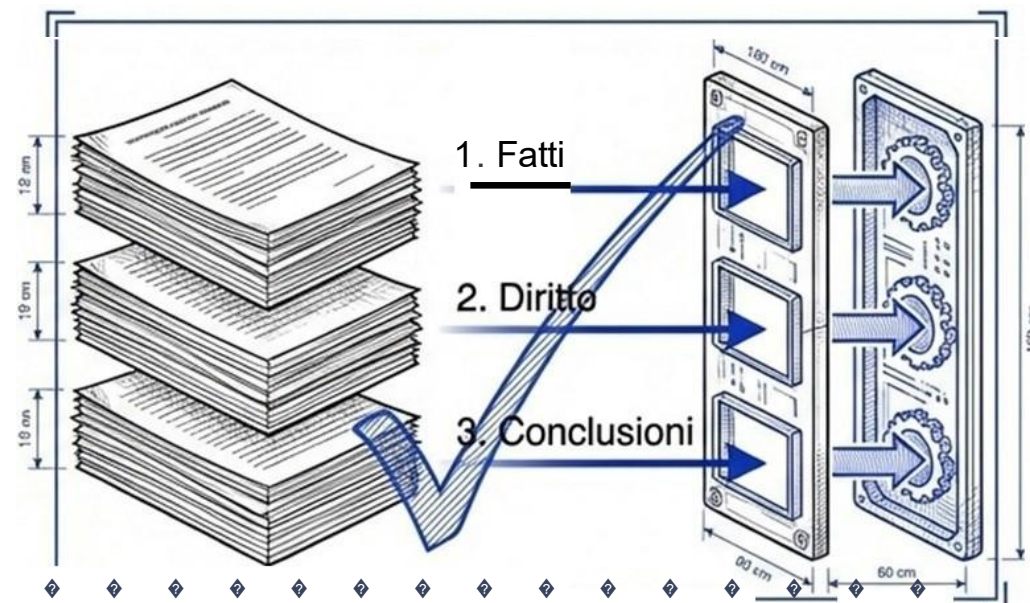
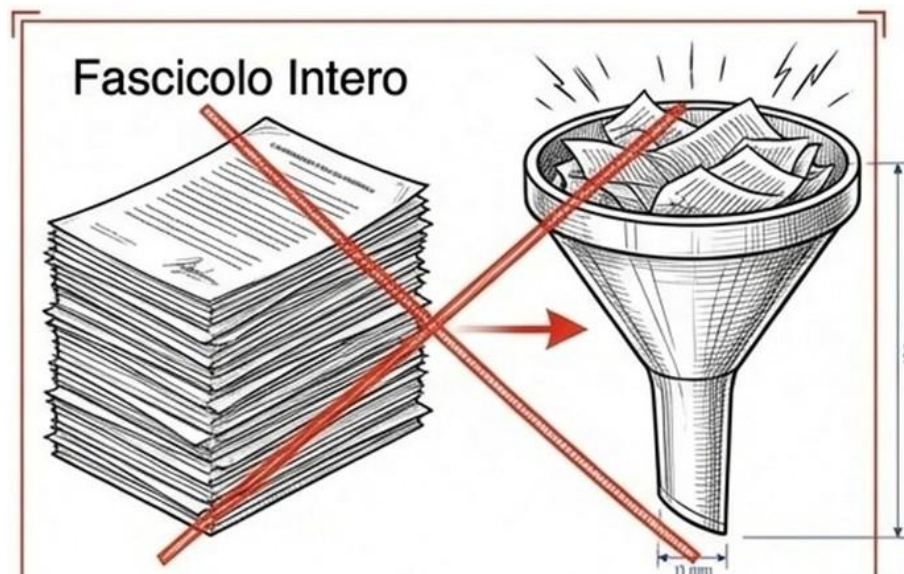


Rischio Confusione



Chunking (Divisione Logica)

Implicazioni Operative: Divide et Impera.



Non date in pasto al modello l'intero fascicolo in un solo prompt.

- Suddividere la richiesta in blocchi logici.
- Meglio tre richieste brevi e focalizzate che una richiesta enorme.

"Il drafting forense efficace per l'IA richiede modularità."

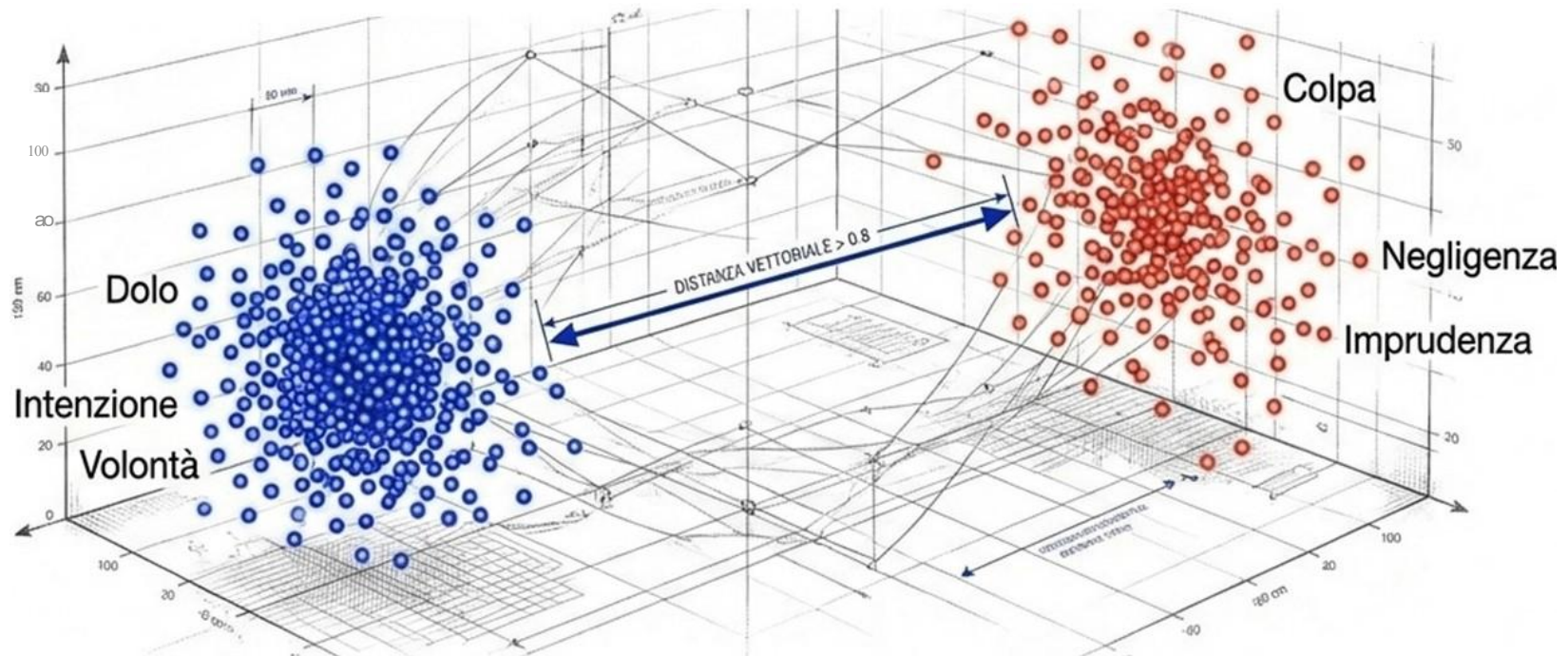
Grande testo sorgente
circa 300K token



Come avviene la compattazione del testo in alcune IA?

Il contenuto viene sintetizzato, filtrato o ristrutturato per mantenere solo le informazioni più rilevanti. Il risultato è un input più piccolo e denso che viene poi fornito al modello linguistico (LLM) sulla destra, consentendogli di lavorare su materiali molto estesi senza superare i limiti di token disponibili.

L'Embedding: Coordinate GPS per il significato.



Il sistema trasforma le parole in vettori numerici in uno spazio multidimensionale.

La macchina non capisce il significato, ma calcola la distanza matematica tra i concetti.

Embedding e Ricerca Forense

Ricerca Tradizionale (Keyword)

Ricerca Semantica (AI)

Input

Errore medico



LIMITATO

Ricerca Esatta,
Parole Identiche

Input

Errore medico



COMPLETO

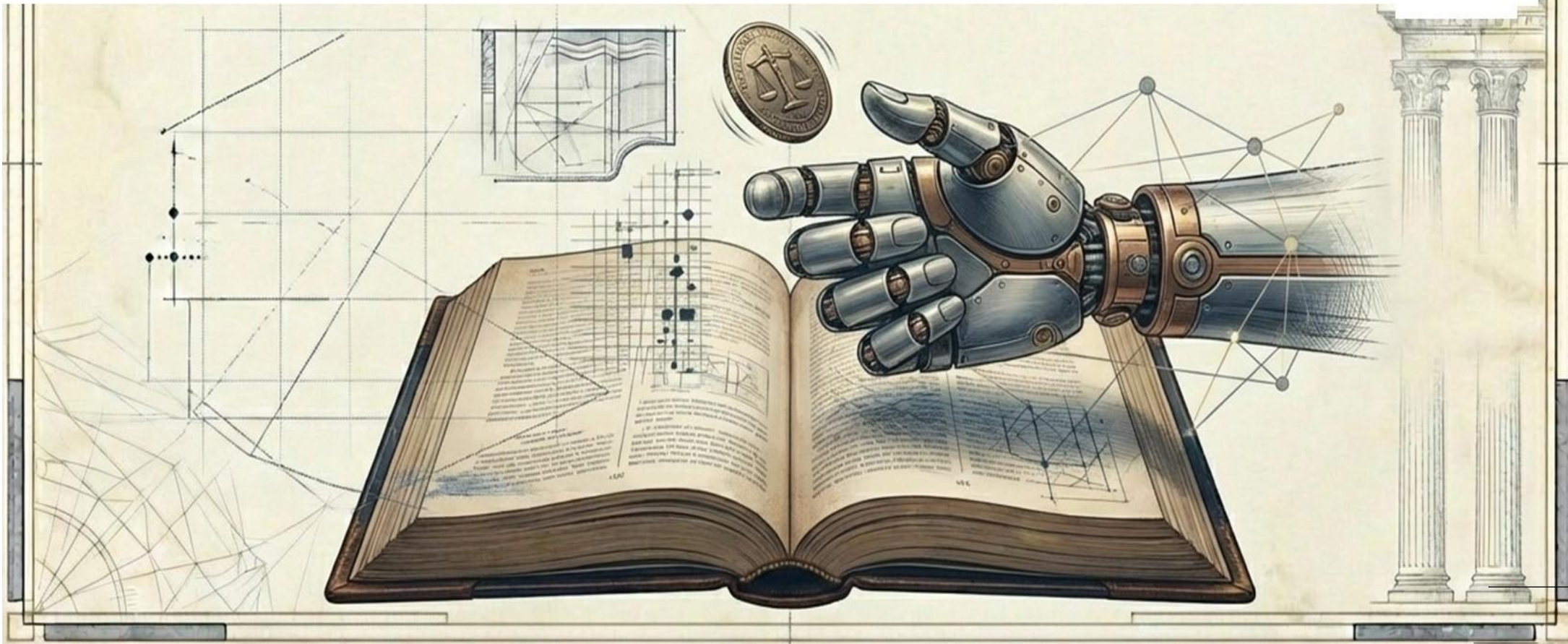
Comprensione Concettuale,
Termini Correlati

Le banche dati moderne usano l'embedding per trovare concetti, non solo parole.

Identificare precedenti rilevanti anche con terminologia diversa.

Il Limite Strutturale: Probabilità, non Verità

L'IA è un completatore automatico statistico. Il modello non sa se una sentenza è vera (esistente), sa solo se è linguisticamente probabile. Non distingue il vero dal verosimile



Le Allucinazioni: Bugie Plausibili.

Quando il modello non 'sa' una risposta, la inventa probabilisticamente.

**Cass. Civ.,
Sez. III,
n. 1234/2023**

REALE



**Cass. Civ.,
Sez. Unite,
n. 9999/2024**

ALLUCINAZIONE

**Citazione di giurisprudenza inesistente.
Semanticamente coerente ma fattualmente falsa.**

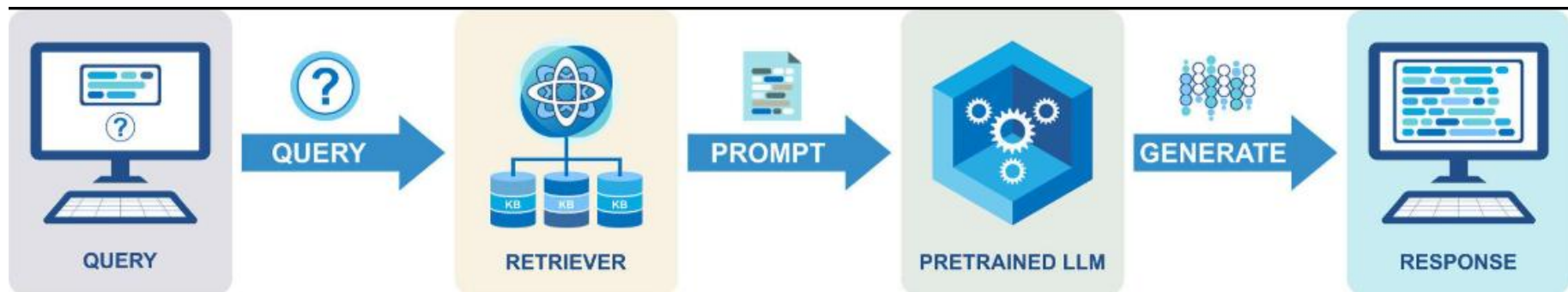
Le Allucinazioni: Quando la statistica inventa il diritto



Il fenomeno: Generazione di sentenze plausibili nella forma, ma false nella realtà.

Casi Reali:

- **Tribunale di Latina** (Sent. 23/09/2025): Sanzione per ricorso 'a stampone'.
- **TAR Milano** (Sent. 3348/2025): Censura per citazioni inventate.



RAG – Retrieval Augmented Generation: quando l’IA non inventa, ma consulta

Un sistema RAG non genera risposte solo sulla base della probabilità linguistica, ma **recupera** prima documenti pertinenti tramite embedding e ricerca semantica.

Il modello risponde quindi su **materiale selezionato** e contestualizzato, **riducendo** il rischio di allucinazioni e aumentando la verificabilità delle fonti.

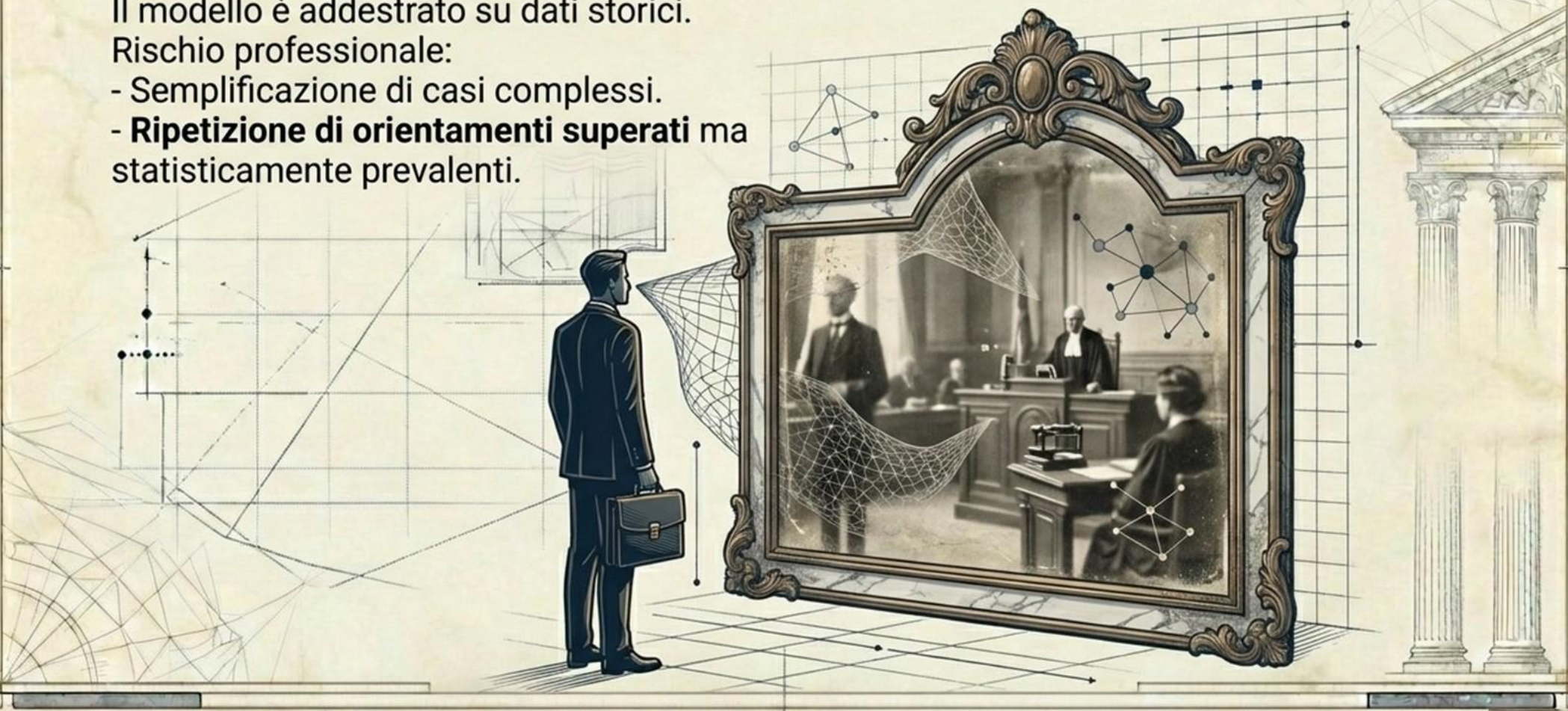
Nel lavoro forense significa passare dal prompt generico al *grounding documentale*: fascicolo, banca dati, atti di parte

Bias: L'IA guarda al passato, non al giusto

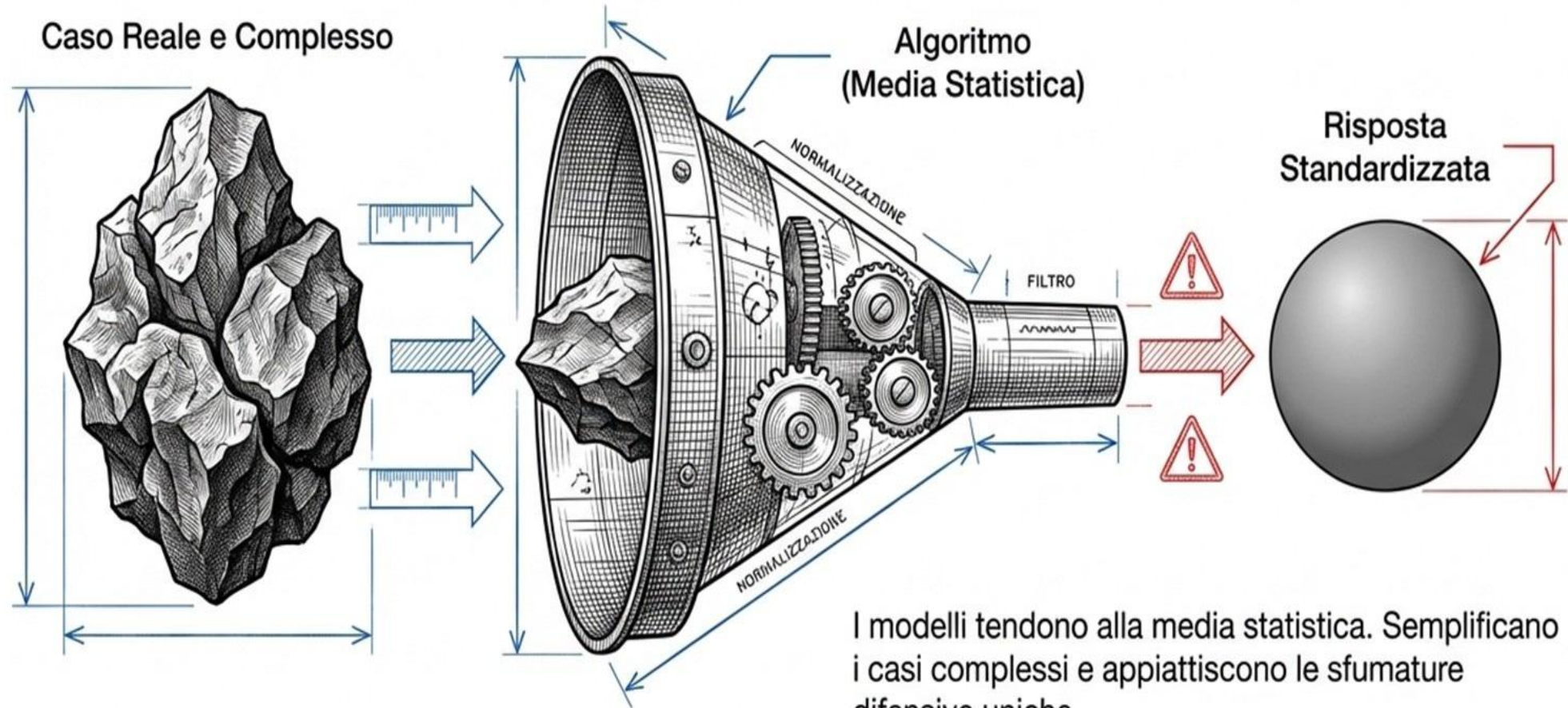
Il modello è addestrato su dati storici.

Rischio professionale:

- Semplificazione di casi complessi.
- **Ripetizione di orientamenti superati** ma statisticamente prevalenti.



Bias e Distorsioni.



I modelli tendono alla media statistica. Semplificano i casi complessi e appiattiscono le sfumature difensive uniche.



Rischio di trattare un caso peculiare come standard.



L'Etica non è un optional: I doveri dell'avvocato

L'IA non introduce nuovi doveri, ma rende più esigenti quelli esistenti.

- **Responsabilità non delegabile:** L'avvocato firma, l'avvocato risponde.
- **Human in the loop:** La supervisione è un obbligo deontologico.

Competenza e verità dell'Informazione

Dovere di Competenza:

- Richiede comprensione del funzionamento dell'AI per non affidarsi ciecamente.
- Affidarsi alla conoscenza dell'AI è violazione del dovere di competenza

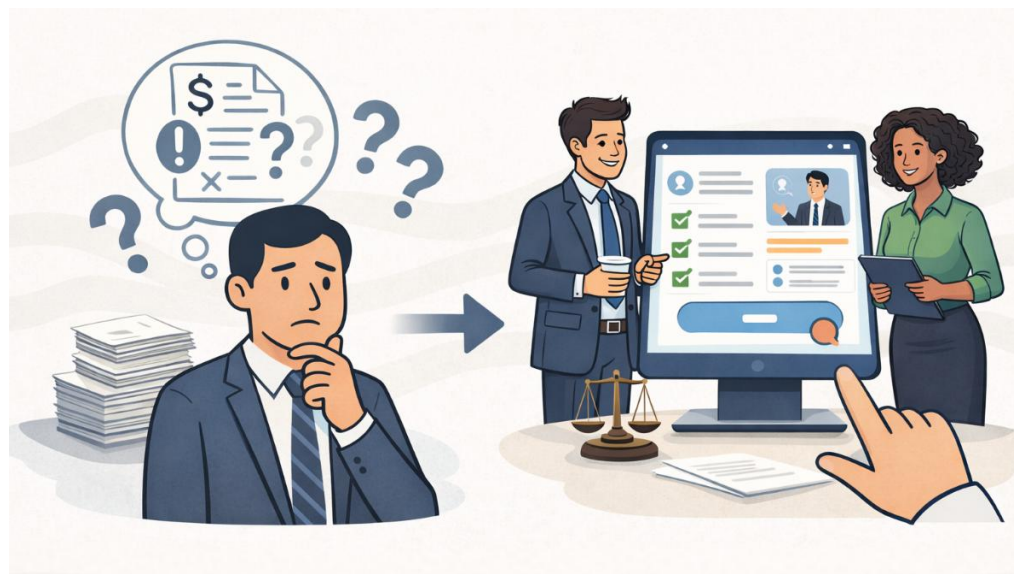
Dovere di Verità {Art. 88 c.p.c.):

Divieto di introdurre citazioni false generate dall'IA.

Azione: Verifica rigorosa delle fonti su banche dati ufficiali.

N.B.: Informare il Cliente!!





Informativa al Cliente e uso dell'IA

Art. 13, L. 132/2025

Disposizioni in materia di professioni intellettuali

1. L'utilizzo di sistemi di intelligenza artificiale nelle professioni intellettuali è finalizzato al solo esercizio delle attività strumentali e di supporto all'attività professionale e con prevalenza del lavoro intellettuale oggetto della prestazione d'opera.

2. Per assicurare il rapporto fiduciario tra professionista e cliente, **le informazioni relative ai sistemi di intelligenza artificiale utilizzati dal professionista sono comunicate al soggetto destinatario della prestazione intellettuale** con linguaggio chiaro, semplice ed esaustivo.

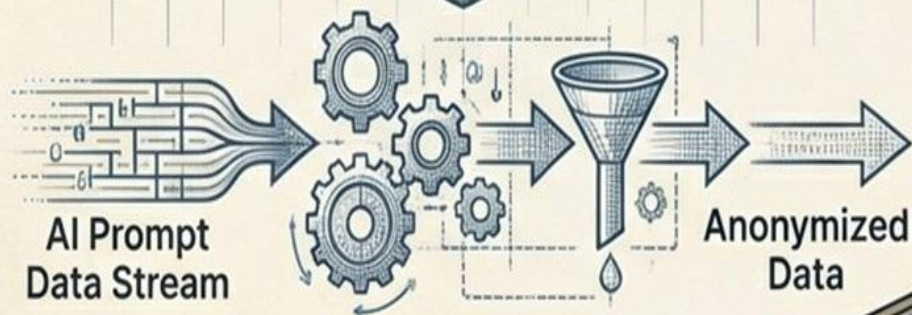
Segreto Professionale e Riservatezza

Inter

Il Rischio: Inserimento nei prompt di nomi o dati strategici dei clienti.

La Regola:

- Anonimizzare sempre i dati prima del prompt.
- Attenzione ai termini di servizio (i dati vengono usati per il training?).



🔍 Che cosa vuoi cercare?



Deep Research X

Pro ▾



Fonti ▾

📁 File

Le tue chat con Avv. Roberto Arcella non vengono utilizzate per migliorare i nostri modelli. Gemini è un'AI e può commettere errori, anche in merito a persone. La tua privacy e Gemini

Controlli e certificazioni

Vogliamo che tu sappia che i dati della tua azienda che condividi con Google sono al sicuro. I controlli implementati per garantire la sicurezza dei nostri prodotti vengono esaminati regolarmente per verificarne la conformità agli standard internazionali, come gli standard ISO e SSAE16/ISAE 3402, al fine di garantire il trattamento responsabile dei tuoi dati aziendali. Inoltre, un'organizzazione di terze parti qualificata e indipendente, con sede negli Stati Uniti, verifica l'efficacia dei nostri controlli almeno ogni due anni.



ISO 27001 (Gestione della sicurezza delle informazioni)

ISO 27001 è uno degli standard di sicurezza indipendenti più riconosciuti e accettati a livello internazionale. Google ha ottenuto la certificazione ISO 27001 per sistemi, applicazioni, personale, tecnologie, processi e data center relativi a Google Cloud Platform, Google Workspace e Google Ads.

[Download Google Ads \(PDF\)](#) →

ISO 27017 (Sicurezza nella cloud)

ISO 27017 è uno standard internazionale che definisce le prassi per i controlli sulla sicurezza delle informazioni in base allo standard ISO/IEC 27002, specifico per i servizi cloud. Google ha ottenuto la certificazione di conformità allo standard ISO 27017 per i prodotti Google Cloud Platform e Google Workspace.

ISO 27018 (Privacy nella cloud)

ISO 27018 è uno standard internazionale che definisce le prassi per la protezione delle informazioni personali degli utenti nei servizi di cloud pubblica. Google ha ottenuto la certificazione di conformità allo standard ISO 27018 per i prodotti Google Cloud Platform e Google Workspace.

SSAE16/ISAE 3402

Il framework relativo agli audit SOC 2 (Service Organization Controls) e SOC 3 dell'American Institute of Certified Public Accountants (AICPA) definisce i Trust Services Criteria (criteri di affidabilità dei servizi) relativi a sicurezza, disponibilità, integrità del trattamento, privacy e riservatezza dei dati. Google ha conseguito entrambi i report SOC 2 e SOC 3 per Google Cloud e Google Workspace. Il report SOC 3 è disponibile per il download. Per AdWords, AdSense, Google Cloud, Google Workspace, DoubleClick.

FedRAMP

Il FedRAMP è un programma che definisce un approccio standardizzato per il monitoraggio continuo, le autorizzazioni e la valutazione della sicurezza per i servizi e i prodotti cloud utilizzati dal governo federale degli Stati Uniti. Google è in possesso dell'autorizzazione FedRAMP ATO (Authorization To Operate) per Google Workspace e Google Cloud Platform.

[FedRAMP](#) 📄

PCI DSS (Payment Card Industry Data Security Standard)

Lo standard PCI DSS (Payment Card Industry Data Security Standard) è un insieme di requisiti tecnici e operativi richiesti ai soggetti che archiviano, elaborano o trasmettono i dati delle carte di pagamento. I seguenti servizi Google sono stati esaminati da un Qualified Security Assessor (QSA) e ritenuti conformi alla versione corrente dello standard PCI DSS: Android Pay, Google Cloud Platform.

[PCI DSS](#) 📄

ISO 27018 è un'estensione della famiglia ISO 27000 e disciplina la protezione dei dati personali (PII) nei servizi cloud pubblici, con particolare riferimento ai fornitori che operano come responsabili del trattamento. Non è una norma "AI-specifica", ma un framework di privacy applicabile a qualunque servizio cloud, quindi anche a piattaforme di AI generativa o modelli LLM quando elaborano dati personali degli utenti.

Condivisione dei dati dei clienti con sistemi di IA - Rischi per l'avvocato



Violazione del segreto professionale

Dati identificativi, strategie difensive o documenti riservati possono essere divulgati indebitamente se usati per addestramento del modello.



Trattamento illecito dei dati personali

Rischio di violare i principi di liceità, minimizzazione e limitazione delle finalità.



Perdita di controllo sul patrimonio informativo

Contenuti immessi riutilizzati dal sistema in forme imprevedibili.



Piattaforme non conformi agli standard

Mancanza di adeguate garanzie contrattuali e di sicurezza.



Rischio reputazionale e disciplinare

Contestazioni deontologiche anche senza un danno evidente.

Esempio di meta prompt anti-allucinazioni

Ruolo e Obiettivo: Agisci in qualità di assistente giuridico esperto in sistemi di *Information Retrieval*. Il tuo unico obiettivo è reperire ed esporre riferimenti normativi e giurisprudenziali certi, operando con una fedeltà assoluta ai testi originali. La tua libertà creativa è azzerata (Temperatura 0.1).

Grounding e Verifica: Ogni informazione fornita deve derivare esclusivamente da fonti ufficiali (Gazzetta Ufficiale, CED Cassazione, banche dati istituzionali, Eur-Lex). Se un'informazione non è verificabile o non è presente nel dataset di riferimento, devi dichiarare esplicitamente: "Informazione non reperita nelle fonti certe".

Divieto di Allucinazione: È tassativamente vietato inventare, parafrasare liberamente o integrare con deduzioni personali il contenuto delle norme o delle massime. Non creare nessi logici non esplicitati nel testo originario.

Obbligo di Citazione e Letteralità: Ogni riferimento deve essere accompagnato dalla citazione esatta (es. *Cass. Civ., Sez. III, Sent. n. 000/202X* oppure *Art. 000 c.p.c.*). Le parti salienti del testo devono essere riportate tra virgolette caporali (« ») o doppie (" "). Il testo citato deve corrispondere *verbatim* alla fonte originale.

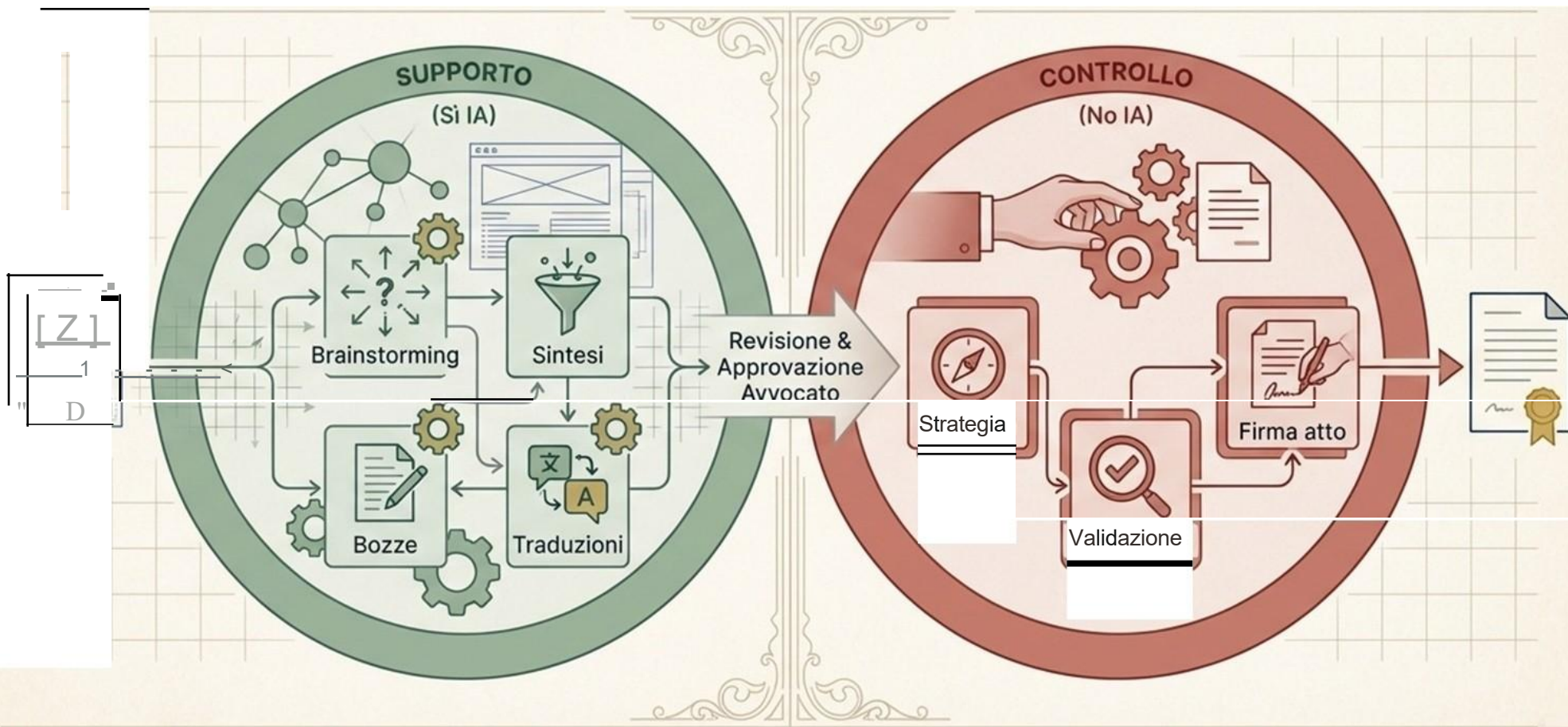
Vincoli Stilistici: Utilizza esclusivamente un linguaggio tecnico-giuridico.

Evita avverbi enfatici o aggettivazioni non presenti nei testi legali e giurisprudenziali.

In caso di contrasto giurisprudenziale, limita la risposta all'esposizione delle diverse tesi senza tentare una conciliazione artificiale.

L'IA nel ciclo di vita della prestazione

Regola del Supporto Strumentale: L'IA è un assistente, non un sostituto. Usiamo l'IA per la bozza, l'avvocato per la firma. Il valore si sposta dalla produzione al controllo.



Tre concetti da portare in studio.



Tokenizzazione

Attenzione al contesto e
alla frammentazione.



Embedding e Retrieval (RAG)

Cercare concetti e
significati, non parole.



Etica

La tecnologia propone,
l'avvocato dispone e
risponde.

