GIUSTIZIA INSIEME

Diritto e Innovazione

IA e machina sapiens della giustizia

Vademecum operativo per l'uso dell'intelligenza artificiale negli uffici giudiziari

di Fernanda Iannone

19 giugno 2026

ABSTRACT

Il presente vademecum traduce in chiave operativa le Raccomandazioni adottate dal Consiglio Superiore della Magistratura con delibera plenaria dell'8 ottobre 2025 sull'uso dell'intelligenza artificiale nell'amministrazione della giustizia. Il documento si rivolge a magistrati, funzionari e personale amministrativo degli uffici giudiziari e si propone di fornire una guida concreta per orientarsi nella fase transitoria che precede la piena operatività del Regolamento (UE) 2024/1689 (AI Act), attesa entro l'agosto 2026. Dopo aver delineato la mappa dei rischi specifici dell'IA nel contesto giudiziario – dalle allucinazioni alle citazioni apocrife, dal bias algoritmico alla prompt injection, dall'automation bias allo slittamento di funzione – il vademecum enuncia i principi fondamentali che governano ogni utilizzo (sovranità dei dati, minimizzazione, verificabilità, responsabilità individuale integra) e traccia con

precisione il perimetro delle attività ammissibili e dei divieti assoluti. La parte finale è dedicata alle istruzioni operative di prompting per Microsoft Copilot e sistemi analoghi, articolate per fasi di lavoro: configurazione del meta-prompt, primo esame degli atti, ricerca giurisprudenziale, supporto alla redazione. Una checklist operativa, un modello di annotazione per la tracciabilità e una tavola sinottica di prompt consentiti e vietati completano lo strumento, pensato per un utilizzo quotidiano immediato e conforme.

Sommario: 1. **Premessa istituzionale** – 2. **Definizioni essenziali** – 3. **La mappa dei rischi: conoscerli per governarli** – 3.1. **Sovranità e sicurezza dei dati** – 3.2. **Allucinazioni e citazioni apocrife** – 3.3. **Risposte sycophantic e bias algoritmici** – 3.4. **Prompt injection** – 3.5. **Automation bias e slittamento di funzione** – 3.6. **Model drift e variabilità degli output** – 4. **I principi fondamentali che governano ogni utilizzo** – 5. **Attività ammissibili: il perimetro del consentito** – 5.1. **Attività ammissibili a titolo esemplificativo** – 5.2. **La distinzione cruciale: sintesi estrattiva e sintesi astrattiva** – 6. **Divieti assoluti: il perimetro dell'inammissibile** – 6.1. **Esempi operativi: prompt consentiti e prompt vietati** – 7. **Ricerche giurisprudenziali assistite: confini e cautele** – 8. **Protezione dei dati, riservatezza e prevenzione della profilazione** – 9. **Sicurezza informatica e gestione degli incidenti** – 10. **Tracciabilità, annotazione e accountability** – 11. **Governance e responsabilità organizzativa** – 12. **Formazione come dovere professionale** – 13. **Checklist operativa preliminare a ogni utilizzo** – 14. **Monitoraggio, audit e aggiornamento del presente vademecum** – 15. **Considerazioni conclusive** – 16. **Istruzioni operative per l'uso di Copilot e sistemi analoghi: tecniche di prompting** – 16.1. **Il meta-prompt di sistema: configurare il modello prima di usarlo** – 16.2. **Primo esame degli atti: organizzare senza valutare** – 16.3. **Ricerca giurisprudenziale e studio del diritto: mappare senza selezionare** – 16.4. **Supporto alla redazione: strutturare senza decidere** – 16.5. **Regole di condotta generale nel prompting**

1. Premessa istituzionale

Con la delibera plenaria dell'8 ottobre 2025 il Consiglio Superiore della Magistratura ha adottato le Raccomandazioni sull'uso dell'intelligenza artificiale nell'amministrazione della giustizia, definendo per la prima volta a livello nazionale un quadro organico di principi, cautele e divieti destinato a orientare l'utilizzo delle tecnologie di intelligenza artificiale da parte di quanti operano negli uffici giudiziari.

Il presente vademecum ha lo scopo di tradurre tali Raccomandazioni in uno strumento di orientamento pratico, immediatamente fruibile da magistrati, funzionari e personale amministrativo nell'attività quotidiana. Non si tratta di un documento puramente teorico né di un catalogo di adempimenti formali: esso intende piuttosto essere una guida concreta, che consenta a ciascun operatore di riconoscere le situazioni in cui l'utilizzo dell'intelligenza artificiale è consentito, quelle in cui è precluso, e di gestire in modo corretto e tracciabile ogni impiego effettuato.

Il vademecum si colloca in una fase transitoria caratterizzata da profonda incertezza tecnologica e rapida evoluzione normativa. Il Regolamento (UE) 2024/1689 – il cosiddetto AI Act – ha qualificato come sistemi ad alto rischio quelli destinati ad assistere l'autorità giudiziaria nella ricerca e nell'interpretazione dei fatti e del diritto nonché nell'applicazione della legge a casi concreti. La piena operatività delle norme di conformità europee è attesa entro l'agosto del 2026; sino a tale scadenza vigono criteri di cautela rafforzata, che questo documento si propone di rendere operativi in modo semplice ed efficace.

L'intelligenza artificiale può e deve essere un alleato degli uffici giudiziari per liberare tempo, ridurre la pressione su attività routinarie e migliorare l'organizzazione del lavoro. Ma essa non può, in alcun caso, sostituirsi al ragionamento del magistrato, alla valutazione delle prove, all'interpretazione della legge, all'adozione del provvedimento giurisdizionale. La centralità della decisione umana non è una resistenza conservatrice all'innovazione: è la garanzia costituzionale che presiede all'esercizio della funzione giurisdizionale in uno Stato di diritto.

La responsabilità individuale dell'utilizzatore rimane integra a prescindere dal comportamento del sistema. Nessun malfunzionamento del programma, nessun output erroneo, nessuna allucinazione del modello può essere invocato a discarico: chi utilizza uno strumento di intelligenza artificiale risponde integralmente delle conseguenze di quell'utilizzo.

2. Definizioni essenziali

Sistema di intelligenza artificiale: qualsiasi sistema automatizzato che, elaborando input di varia natura, genera output – testi, raccomandazioni, previsioni, decisioni – con livelli variabili di autonomia rispetto all'intervento umano.

Intelligenza artificiale generativa: la categoria di sistemi capaci di produrre contenuti nuovi – testi, immagini, codice – a partire da istruzioni testuali denominate prompt. In questa categoria rientrano i più diffusi chatbot commerciali attualmente liberamente accessibili in rete.

Sistema di IA autorizzato: il sistema messo a disposizione o espressamente abilitato dal Ministero della Giustizia nell'ambito del dominio istituzionale della giustizia, ovvero autorizzato ai sensi della normativa nazionale vigente e delle policy interne dell'amministrazione.

Sistema di IA generalista o di uso generale: il sistema destinato a impieghi generici e diffusi presso il pubblico, non specificamente abilitato per il dominio della giustizia. L'utilizzo di tali strumenti in connessione con l'attività giudiziaria è soggetto alle restrizioni più severe illustrate nel presente vademecum.

Attività giudiziaria in senso stretto: l'attività funzionale alla decisione, che comprende l'interpretazione e l'applicazione della legge, la valutazione dei fatti e delle prove, nonché l'adozione del provvedimento giurisdizionale. È questo il nucleo insostituibile della funzione giurisdizionale, sul quale si concentrano i divieti più incisivi.

Dati riservati: i dati personali di qualsiasi categoria, i dati giudiziari, i segreti d'ufficio, gli atti non ostensibili, le informazioni coperte da segreto investigativo o istruttorio, nonché ogni elemento idoneo a consentire, anche indirettamente, la reidentificazione di persone fisiche o di fascicoli processuali.

3. La mappa dei rischi: conoscerli per governarli

Una corretta cultura dell'intelligenza artificiale inizia dalla conoscenza dei rischi specifici che i sistemi in questione comportano nel contesto giudiziario. Di seguito si illustrano i principali, con l'indicazione delle cautele operative conseguenti.

3.1. Sovranità e sicurezza dei dati

Ogni informazione inserita in un sistema di intelligenza artificiale esterno al dominio istituzionale viene trasmessa a server che possono essere collocati al di fuori dello Spazio economico europeo, dove non trovano applicazione le garanzie del Regolamento generale sulla

protezione dei dati. Il materiale immesso può essere trattenuto, riutilizzato per addestrare versioni future del modello o divenire accessibile a soggetti terzi non autorizzati. È necessario tenere presente, in particolare, che anche dati formalmente anonimizzati o pseudonimizzati possono essere riidentificati attraverso tecniche di correlazione con altri dataset: il rischio non si esaurisce nella singola informazione trasmessa, ma si estende alle inferenze che il sistema è in grado di elaborare dall'insieme degli input ricevuti.

3.2. Allucinazioni e citazioni apocrife

I modelli di linguaggio generativi non consultano archivi verificati: essi producono testi statisticamente plausibili sulla base di pattern appresi durante l'addestramento, indipendentemente dalla loro corrispondenza alla realtà. Il fenomeno delle cosiddette allucinazioni – generazione di contenuti privi di fondamento reale – è particolarmente insidioso in ambito giuridico. Il sistema può citare sentenze inesistenti attribuendo loro estremi precisi e apparentemente credibili, attribuire a una norma un contenuto che essa non ha, o inventare riferimenti dottrinali con la stessa fluidità con cui produce testi corretti. Ogni riferimento normativo o giurisprudenziale presente in un output deve essere verificato puntualmente su fonti ufficiali prima di qualsiasi utilizzo: tale verifica non è una formalità, ma la condizione necessaria per l'utilizzabilità dell'output.

3.3. Risposte sycophantic e bias algoritmici

I sistemi generativi tendono a orientare le proprie risposte in direzione di ciò che percepiscono come gradito o atteso dall'interlocutore, indipendentemente dalla correttezza del contenuto. Tale fenomeno, denominato sycophancy, può indurre il sistema a confermare ipotesi erronee formulate nella domanda o a produrre rassicurazioni non fondate. In parallelo, i modelli di intelligenza artificiale incorporano i pregiudizi presenti nei dati di addestramento: possono quindi veicolare bias relativi a genere, provenienza geografica, condizione economica o etnia, in modo non sempre evidente all'utilizzatore. In ambito giudiziario, questo rischio può incidere sull'equità e sull'imparzialità del provvedimento.

3.4. Prompt injection

La prompt injection è una forma di manipolazione che sfrutta la struttura stessa dell'interazione con il modello: mediante istruzioni inserite nell'input – anche in modo non palese, incorporato nel testo oggetto di analisi – si può indurre il sistema a disattendere i propri vincoli operativi o a produrre output che travalicano le finalità consentite. In ambito giudiziario, tale rischio si manifesta quando un atto processuale sottoposto a revisione contenga istruzioni testuali idonee

a orientare il sistema verso valutazioni interpretative, selezioni di precedenti o suggerimenti argomentativi; ovvero quando l'input venga formulato in modo tale da indurre il sistema a trattare come pubblici dati non anonimizzati. La consapevolezza di questo meccanismo deve guidare sia la scelta delle operazioni da affidare al sistema, sia la vigilanza sul contenuto degli atti trasmessi.

3.5. Automation bias e slittamento di funzione

Il rischio più insidioso – proprio perché si sviluppa in modo graduale e inconsapevole – è quello dell'automation bias: la tendenza a conferire progressivamente eccessiva fiducia all'output del sistema, riducendo il controllo critico personale. A questo si affianca il rischio di slittamento di funzione: uno strumento inizialmente adibito a compiti organizzativi o redazionali può essere investito, per inerzia operativa, di compiti valutativi che appartengono esclusivamente alla funzione giurisdizionale. Mantenere una vigilanza attiva e costante su questi fenomeni non è opzionale: è il presupposto perché l'utilizzo dell'intelligenza artificiale resti entro i confini del consentito e non eroda, in modo progressivo e silenzioso, la qualità e l'indipendenza della decisione giudiziaria.

3.6. Model drift e variabilità degli output

I modelli di intelligenza artificiale sono soggetti ad aggiornamenti continui da parte dei loro sviluppatori: il comportamento del sistema può mutare nel tempo anche a parità di input, rendendo difficile la ricostruzione ex post del percorso operativo seguito. Ciò comporta specifiche implicazioni sul piano della tracciabilità e dell'accountability: è necessario annotare non solo lo strumento impiegato, ma anche la versione utilizzata al momento dell'interrogazione.

4. I principi fondamentali che governano ogni utilizzo

Le Raccomandazioni Csm individuano un insieme di principi che costituiscono il quadro di riferimento per ogni decisione operativa sull'utilizzo dell'intelligenza artificiale. Essi devono essere interiorizzati da ciascun operatore non come vincoli burocratici, ma come criteri di orientamento che derivano direttamente dai valori fondamentali dell'ordinamento.

Sovranità dei dati e delle informazioni. Le informazioni conferite ai sistemi di intelligenza artificiale e gli output da essi generati non possono essere resi accessibili a soggetti terzi non autorizzati né trasferiti al di fuori dell'area di controllo dell'amministrazione. La tutela della sovranità informativa è il presupposto di ogni utilizzo legittimo.

Minimizzazione e proporzionalità dell'input. Si inserisce nei sistemi di intelligenza artificiale soltanto ciò che è strettamente necessario al conseguimento dell'obiettivo dichiarato. Gli input vengono sempre ridotti alla quantità minima indispensabile, privilegiando sistematicamente la forma anonimizzata, e si tiene debito conto del rischio di reidentificazione anche ove i dati siano stati previamente trattati.

Finalità determinata e divieto di riutilizzo. L'intelligenza artificiale è impiegata esclusivamente per la finalità dichiarata e consentita. L'output generato non può essere riutilizzato per scopi ulteriori non previamente valutati, anche qualora esso risulti formalmente idoneo ad altri impieghi.

Qualità dei dati e non discriminazione. L'operatore verifica la rappresentatività e l'affidabilità degli input e non dà seguito a output che si rivelino affetti da bias evidenti o da contenuti discriminatori. L'utilizzo di dati di addestramento non rappresentativi può produrre output sistematicamente distorti, con ricadute sull'equità della funzione.

Verificabilità delle fonti e riscontro su repertori ufficiali. Ogni riferimento normativo o giurisprudenziale contenuto nell'output è oggetto di puntuale riscontro su fonti ufficiali e autorevoli prima di qualsiasi utilizzo. Il rischio di citazioni inesistenti, imprecise o apocrife è intrinseco ai sistemi generativi e non può essere eliminato: può soltanto essere governato attraverso la verifica sistematica.

Supervisione e replicabilità. L'operatore è in grado di replicare autonomamente le conclusioni attraverso fonti certificate, indipendentemente dall'output del sistema. I risultati che si rivelino inattendibili, parziali o non verificabili sono corretti o scartati. L'output del sistema non costituisce mai una fonte autonoma di conoscenza giuridica.

Responsabilità individuale integra e non delegabile. L'utilizzatore conserva la piena responsabilità per ogni impiego effettuato. Nessun malfunzionamento del sistema, nessun output erroneo, nessuna lacuna del modello può essere invocato a discolpa. La responsabilità non si trasferisce né al sistema né alla struttura di governance: essa permane integralmente in capo alla persona fisica che ha effettuato l'utilizzo.

Centralità assoluta del pensiero critico. L'intelligenza artificiale non può in alcun caso sostituirsi al ragionamento del magistrato né comprimere l'originalità della riflessione giuridica. La vigilanza costante contro l'automation bias – la tendenza a ridurre il controllo critico in favore dell'output del sistema – è un imperativo funzionale che discende direttamente

dall'ordinamento costituzionale e che deve essere coltivato attivamente, non presunto come acquisito.

5. Attività ammissibili: il perimetro del consentito

Nella fase transitoria vigente, le Raccomandazioni Csm individuano nella clausola di cui all'articolo 6, paragrafo 3, del Regolamento (UE) 2024/1689 lo spazio operativo entro cui è possibile muoversi. Tale norma consente l'utilizzo di sistemi di intelligenza artificiale, ancorché non ancora conformi ai requisiti dell'AI Act per i sistemi ad alto rischio, per compiti di natura procedurale, organizzativa o accessoria, che non esercitino un'influenza materiale sull'esito del processo decisionale.

Le condizioni che devono concorrere affinché l'utilizzo sia legittimo ai sensi di tale deroga sono le seguenti: l'impiego avviene esclusivamente nell'ambito degli applicativi del dominio istituzionale della giustizia o di sistemi espressamente autorizzati dal Ministero; l'utilizzo è effettuato in modalità tracciata e sicura; è prevista e attuata una revisione umana sistematica e completa dell'output prodotto.

5.1. Attività ammissibili a titolo esemplificativo

Rientrano nel perimetro delle attività consentite, a titolo esemplificativo e non esaustivo, le seguenti operazioni:

- Elaborazione di schemi organizzativi, prospetti e tabelle per la gestione delle udienze, dei ruoli e dei carichi di lavoro.
- Redazione di bozze di comunicazioni, verbali e documenti di carattere amministrativo-organizzativo privi di contenuto decisionale.
- Ottimizzazione della calendarizzazione, dei turni e delle assegnazioni sulla base di criteri predeterminati.
- Supporto alla stesura di note, relazioni statistiche, comunicazioni interne e materiali destinati alla formazione.
- Comparazione di prassi organizzative tra uffici per l'individuazione di criticità e buone pratiche replicabili.
- Generazione di presentazioni, tabelle e grafici su dati aggregati e non identificativi.
- Revisione linguistica e stilistica di testi sotto il profilo della coerenza sintattica, dell'ortografia e della chiarezza espositiva, senza alterazione del contenuto sostanziale.
- Catalogazione e archiviazione per materia dei quesiti ai consulenti tecnici d'ufficio secondo parole chiave o categorie tematiche.

- Traduzione assistita di testi non riservati o previamente anonimizzati, con successiva verifica integrale da parte dell'operatore.

5.2. La distinzione cruciale: sintesi estrattiva e sintesi astrattiva

Un elemento di particolare rilevanza operativa riguarda la distinzione tra la sintesi estrattiva e la sintesi astrattiva degli atti processuali. La distinzione non dipende dalla lunghezza del testo prodotto né dalla complessità apparente dell'operazione, bensì dalla presenza o assenza di un criterio selettivo autonomo esercitato dal sistema.

La sintesi estrattiva consiste nell'individuazione e riproduzione di parti testuali già presenti nell'atto, senza alcuna rielaborazione concettuale e senza alcuna valutazione di rilevanza. In tale caso il sistema non compie alcuna operazione valutativa: si limita a isolare e riprodurre ciò che è già espresso nel testo. Tale operazione è ammessa.

La sintesi astrattiva comporta invece una selezione autonoma dei fatti o delle questioni ritenute rilevanti, la riformulazione delle argomentazioni, la gerarchizzazione degli elementi in funzione di un giudizio di importanza. In tale caso il sistema esercita una valutazione che appartiene esclusivamente alla sfera del ragionamento giuridico del magistrato. Tale operazione non è ammessa.

Quando il prompt chiede al sistema di individuare ciò che è «rilevante», «decisivo», «importante», «convincente» o «fondamentale» ai fini del giudizio, l'operazione cessa di essere estrattiva e diventa astrattiva: il sistema non riproduce, ma valuta. Tale operazione è preclusa.

6. Divieti assoluti: il perimetro dell'inammissibile

I divieti che seguono sono tassativi e non suscettibili di interpretazione estensiva a favore dell'utilizzo. Il loro fondamento non è meramente tecnico, ma assiologico: essi presidiano l'insostituibilità della decisione umana nell'esercizio della funzione giurisdizionale e discendono direttamente dai principi costituzionali che ne presiedono l'esercizio.

È fatto divieto assoluto di utilizzare sistemi di intelligenza artificiale non autorizzati nell'ambito dell'attività giudiziaria in senso stretto, comprensiva della valutazione dei fatti e delle prove, dell'interpretazione e dell'applicazione della legge al caso concreto, e dell'adozione del provvedimento giurisdizionale.

È fatto divieto assoluto di immettere in sistemi di intelligenza artificiale generalista o in strumenti esterni al dominio istituzionale atti processuali non ostensibili o dati riservati di qualsiasi categoria, anche qualora tali informazioni siano fornite in forma indiretta o potenzialmente reidentificabile.

Non è consentito delegare all'intelligenza artificiale valutazioni attinenti alla credibilità delle fonti di prova, alla ponderazione degli elementi probatori, alla determinazione della pena o della sanzione applicabile, né richiedere al sistema proposte di motivazione fondate su elaborazioni algoritmiche.

È vietata qualsiasi forma di profilazione o classificazione di persone fisiche – parti processuali, indagati, imputati, testimoni, vittime o soggetti minorenni – mediante sistemi di intelligenza artificiale.

Rimane precluso l'utilizzo dell'intelligenza artificiale al fine di eludere le regole di accesso alle banche dati ovvero di aggirare le misure di sicurezza informatica adottate dall'amministrazione.

6.1. Esempi operativi: prompt consentiti e prompt vietati

Prompt consentiti (su testo anonimizzato)	Prompt vietati
«Riorganizza il presente testo in punti chiari, astenendoti dall'aggiungere fatti nuovi e dal formulare conclusioni giuridiche.»	«Suggerisci quale decisione adottare nel presente caso e come articolarne la motivazione.»
«Individua i refusi e le ripetizioni presenti nel testo, suggerendo miglioramenti stilistici che lascino invariato il contenuto sostanziale.»	«Valuta la credibilità del testimone sulla base delle sue dichiarazioni e indica se esse ti appaiono coerenti.»
«Estrai dall'atto l'indicazione delle parti, le conclusioni rassegnate e l'elenco dei motivi così come numerati e rubricati, senza rielaborarli.»	«Stima la pena congrua tenendo conto delle circostanze attenuanti e aggravanti descritte nell'atto.»
«Elenca i riferimenti normativi e giurisprudenziali espressamente citati nell'atto, senza aggiungere commenti o valutazioni.»	«Sintetizza l'atto evidenziando i fatti principali e le questioni giuridiche rilevanti ai fini della decisione.»
«Elabora una checklist dei passaggi formali da verificare in un provvedimento di questo tipo, senza suggerire alcun esito.»	«Individua le debolezze degli argomenti difensivi e indica quale linea argomentativa ritieni più convincente.»

Prompt consentiti (su testo anonimizzato)	Prompt vietati
«Riporta testualmente le conclusioni rassegnate dalle parti, senza rielaborarle.»	«Analizza gli atti allegati e indica la strategia processuale più opportuna da adottare.»

7. Ricerche giurisprudenziali assistite: confini e cautele

Il supporto dell'intelligenza artificiale alle ricerche giurisprudenziali costituisce uno dei casi di utilizzo più frequenti tra gli operatori e, al tempo stesso, uno dei più delicati sotto il profilo della conformità. Le Raccomandazioni Csm ne ammettono l'impiego esclusivamente quale ausilio tecnico-procedurale per la costruzione di stringhe di ricerca, la classificazione tematica e il recupero di materiale documentale, a condizione che il magistrato mantenga in ogni momento un ruolo attivo, critico e non delegabile.

Il confine tra utilizzo consentito e utilizzo non consentito è tracciato dalla presenza o assenza di un criterio selettivo autonomo da parte del sistema. Quando il sistema si limita a restituire materiale sulla base di parole chiave fornite dall'operatore, la funzione è di recupero documentale: tale operazione è ammessa. Quando invece il sistema procede alla selezione automatica della giurisprudenza ritenuta più rilevante, suggerisce orientamenti prevalenti o genera schemi motivazionali, l'utilizzo acquista un'incidenza potenziale sull'attività valutativa che lo rende non consentito nella fase transitoria, salvo che non avvenga nell'ambito di un ambiente espressamente autorizzato o sperimentale.

Le basi dati impiegate devono essere complete, non discriminatorie e costantemente aggiornate, ovvero assoggettate a forme di supervisione nella selezione, classificazione e aggiornamento dei contenuti. È necessario ricordare con fermezza che il rischio di citazioni giurisprudenziali inesistenti o apocrife è intrinsecamente elevato nei sistemi generativi, a prescindere dall'apparenza di credibilità con cui tali citazioni vengono formulate: ogni riferimento deve essere riscontrato puntualmente su repertori ufficiali.

Non ci si può fidare di una sentenza citata dall'intelligenza artificiale sulla sola base della precisione con cui sono indicati gli estremi. I modelli generativi sono in grado di formulare riferimenti giurisprudenziali del tutto inesistenti – inclusi numero di registro, anno, sezione e massima – con la stessa fluidità con cui citano decisioni reali. La verifica su fonti ufficiali è sempre obbligatoria.

8. Protezione dei dati, riservatezza e prevenzione della profilazione

La tutela dei dati personali e delle informazioni riservate costituisce un presidio fondamentale dell'intero impianto. Le Raccomandazioni Csm dedicano particolare attenzione al rischio di profilazione indebita dell'attività giudiziaria e dei singoli magistrati, che può derivare dall'applicazione di tecniche di recupero e rielaborazione di informazioni – in particolare le cosiddette pipeline RAG, ovvero retrieval augmented generation – sugli archivi digitali della giustizia.

Tale fenomeno, tendenzialmente invisibile agli utilizzatori finali, consiste nella possibilità che un sistema addestrato su grandi archivi di provvedimenti sia in grado di inferire stili decisionali, orientamenti interpretativi e pattern di produttività individuale: una forma di sorveglianza dell'attività giudiziaria tecnicamente possibile e altamente lesiva dell'indipendenza della magistratura. Per tale ragione, la creazione di repository o pipeline RAG non autorizzate è espressamente vietata.

La progettazione di soluzioni basate sull'intelligenza artificiale deve privilegiare modelli residenti su server sotto il controllo del Ministero della Giustizia, ovvero modelli anche open source eseguibili in locale su hardware gestito dall'amministrazione, previa adeguata sperimentazione e verifica dei requisiti di sicurezza. Le Raccomandazioni prendono sul punto una posizione netta che merita di essere sottolineata: le garanzie contrattuali concordate con operatori commerciali non sono di per sé sufficienti a soddisfare il livello di tutela richiesto dall'ordinamento. Il ricorso a soluzioni istituzionali interne deve essere trattato come criterio di priorità, non come opzione residuale.

9. Sicurezza informatica e gestione degli incidenti

Sul piano della sicurezza informatica vigono le seguenti prescrizioni inderogabili.

È assolutamente vietato inserire nei prompt o negli strumenti di intelligenza artificiale credenziali di accesso, token di autenticazione, chiavi crittografiche o qualsiasi altra informazione attinente alla sicurezza dei sistemi informatici dell'amministrazione. L'inserimento di tali informazioni espone l'intera infrastruttura a rischi di compromissione che possono avere conseguenze di assoluta gravità.

Gli output prodotti dai sistemi di intelligenza artificiale sono conservati esclusivamente su sistemi autorizzati dall'amministrazione. È precluso l'utilizzo di servizi cloud personali, di account privati o di applicativi non istituzionali per la gestione, archiviazione o trasmissione di materiale d'ufficio.

Qualsiasi anomalia rilevata nell'utilizzo di un sistema di intelligenza artificiale – ivi compresi possibili eventi di data leak, output anomali che rivelino l'esposizione di dati non conferiti, o accessi non autorizzati – deve essere segnalata tempestivamente ai referenti per l'ICT e al presidio competente dell'ufficio. La tempestività della segnalazione è essenziale per l'attivazione delle procedure di incident response e per la limitazione dei danni.

10. Tracciabilità, annotazione e accountability

Le Raccomandazioni Csm introducono un obbligo di tracciamento dell'utilizzo dell'intelligenza artificiale che, pur essendo congegnato in forma semplificata, riveste una funzione essenziale sia sul piano dell'accountability interna sia su quello della eventuale ricostruzione ex post del percorso operativo seguito.

Ciascun ufficio giudiziario adotta una modalità di tracciamento mediante registro interno o annotazione nel fascicolo di lavoro. Il contenuto minimo dell'annotazione comprende i seguenti elementi: la denominazione e la versione dello strumento impiegato; la finalità concretamente perseguita; la tipologia di input, con indicazione espressa del suo carattere anonimizzato o non riservato; il nominativo del soggetto che ha proceduto alla verifica dell'output; l'esito di tale verifica; la destinazione o la modalità di conservazione dell'output.

A tale fine si raccomanda l'adozione, con gli adattamenti del caso, della seguente formula standardizzata:

> «In data _____, per finalità di _____, è stato utilizzato lo strumento _____ (dominio giustizia / autorizzato dal Ministero), versione _____. Sono stati inseriti esclusivamente testi e dati previamente anonimizzati e non riservati. L'output è stato verificato integralmente da _____ con esito _____ (conforme / conforme con correzioni / scartato) e utilizzato esclusivamente come supporto redazionale / organizzativo. L'output è conservato in _____.»

Quanto alla trasparenza verso l'esterno, le Raccomandazioni non impongono in via generale un obbligo di disclosure nei confronti delle parti o del pubblico, in assenza di specifiche disposizioni normative o di policy. Rimane fermo, in ogni caso, il dovere di correttezza e trasparenza processuale nei casi in cui l'utilizzo dell'intelligenza artificiale abbia inciso su profili rilevanti per il procedimento e sussistano specifiche indicazioni istituzionali in tal senso.

11. Governance e responsabilità organizzativa

Le Raccomandazioni Csm delineano un assetto di governance multilivello in cui la responsabilità si articola tra soggetti istituzionali distinti, ciascuno con funzioni specifiche e non sovrapponibili.

La Presidenza dell'ufficio capodistretto adotta l'atto di indirizzo organizzativo, promuove le iniziative formative e le attività di monitoraggio, e trasmette ai competenti tavoli istituzionali le criticità riscontrate e le proposte di miglioramento. I Magistrati referenti per l'informatica supportano l'implementazione nei singoli uffici, raccolgono le esigenze degli operatori, diffondono le buone pratiche individuate e gestiscono il canale di segnalazione dei rischi e degli incidenti. I referenti per l'ICT e per la sicurezza informatica definiscono le misure tecniche e organizzative, le procedure di incident response e i criteri di conservazione degli output. Il Responsabile della protezione dei dati fornisce supporto alle valutazioni in materia di protezione dei dati personali e, ove necessario, coordina le valutazioni d'impatto connesse a eventuali sperimentazioni.

Gli utilizzatori finali – magistrati e personale amministrativo – sono il fulcro dell'intero sistema: a loro spetta l'osservanza concreta delle regole, la verifica sistematica degli output e l'assicurazione di una tracciabilità minima. La governance organizzativa non attenua in alcuna misura la responsabilità individuale di chi effettua l'utilizzo.

12. Formazione come dovere professionale

Le Raccomandazioni Csm assegnano alla formazione continua un ruolo centrale, che non può essere ricondotto a un mero adempimento burocratico. La conoscenza dei rischi, delle regole e degli strumenti di verifica è la precondizione affinché i principi enunciati trovino effettiva applicazione nella pratica quotidiana degli uffici.

I percorsi formativi promossi in collaborazione con la Scuola Superiore della Magistratura e con le strutture competenti sono rivolti a illustrare le potenzialità e i limiti intrinseci dei sistemi di intelligenza artificiale, le tipologie di errore più ricorrenti, i fenomeni di bias algoritmico e model drift, nonché le tecniche di verifica degli output. Particolare rilievo è attribuito ai principi di minimizzazione degli input, al divieto di trattare atti e dati di procedimento, alla triangolazione delle fonti e alla replicabilità dei risultati attraverso fonti ufficiali.

Partecipare attivamente ai percorsi formativi non è soltanto un'opportunità: è una componente del dovere di responsabilità individuale che le Raccomandazioni pongono a carico di ciascun utilizzatore. Chi non partecipa non è nella posizione di valutare correttamente i rischi connessi all'utilizzo che effettua e, quindi, non è nelle condizioni di adempiere alle regole che su di lui incombono.

13. Checklist operativa preliminare a ogni utilizzo

Prima di ricorrere a un sistema di intelligenza artificiale, l'operatore è tenuto a verificare i seguenti elementi. In presenza di una risposta negativa a uno dei primi quattro punti, l'utilizzo non può avere luogo.

Lo strumento rientra nel dominio istituzionale della giustizia oppure è stato espressamente autorizzato dal Ministero della Giustizia?

Si è certi di non inserire dati riservati, elementi identificativi, numeri di procedimento, atti non ostensibili o informazioni coperte da segreto?

L'input è stato ridotto alla quantità minima necessaria e anonimizzato nella misura massima consentita?

L'attività è di natura procedurale, organizzativa o accessoria, senza incidenza sulla decisione giurisdizionale?

È prevista la verifica umana completa dell'output, comprensiva del riscontro su fonti ufficiali?

L'output sarà conservato esclusivamente su sistemi autorizzati dall'amministrazione?

L'utilizzo sarà annotato con le informazioni minime prescritte?

Se tutte le risposte sono affermative, l'utilizzo può procedere. La checklist non esaurisce la responsabilità dell'operatore: essa costituisce il punto di partenza del processo di valutazione, non il suo punto di arrivo.

14. Monitoraggio, audit e aggiornamento del presente vademecum

Il presidio competente svolge attività di monitoraggio periodico, raccoglie le criticità emerse e le proposte elaborate dagli uffici, e trasmette alla Presidenza un report con cadenza semestrale. Il raccordo istituzionale con il Ministero della Giustizia e con il Consiglio Superiore della Magistratura avviene anche attraverso il contributo ai tavoli tecnici e la promozione di ambienti di sperimentazione controllata – le cosiddette sandbox – coerenti con gli indirizzi consiliari.

Il presente vademecum è soggetto ad aggiornamento in caso di nuove indicazioni ministeriali o consiliari, di evoluzioni del quadro normativo europeo e nazionale, ovvero degli esiti di sperimentazioni condotte nell'ambito del dominio istituzionale. L'attività di audit si articola su due livelli distinti: un audit di processo, che verifica il rispetto dei divieti, la corretta gestione dei dati e l'effettività della tracciabilità; un audit di impatto, che valuta gli effetti prodotti sulla qualità degli atti, sui tempi di definizione dei procedimenti, sui profili di bias e sulla sicurezza informatica.

Gli operatori sono incoraggiati a segnalare ai referenti competenti non soltanto le criticità e i malfunzionamenti, ma anche le buone prassi e i casi di utilizzo che si rivelino particolarmente efficaci: la costruzione di un patrimonio condiviso di esperienze positive è uno degli strumenti più efficaci per elevare la qualità complessiva dell'utilizzo.

15. Considerazioni conclusive

L'intelligenza artificiale è entrata negli uffici giudiziari e non ne uscirà: è uno strumento destinato a crescere in pervasività e sofisticazione. La domanda non è se utilizzarla, ma come farlo in modo conforme ai valori fondamentali dell'ordinamento e alle regole che la comunità istituzionale si è data.

Le Raccomandazioni del Consiglio Superiore della Magistratura e il quadro normativo europeo offrono una risposta articolata a tale domanda. Il presente vademecum ha cercato di tradurla in termini operativi, nella convinzione che la conformità non sia il nemico dell'efficienza, ma la sua precondizione: un utilizzo corretto, tracciato e critico dell'intelligenza artificiale produce benefici duraturi; un utilizzo incauto, non verificato e non governato espone l'ufficio, il magistrato e le parti a rischi di cui è difficile prevedere le conseguenze.

Il pensiero critico, la verifica sistematica, la responsabilità individuale non sono ostacoli all'innovazione: sono le condizioni che rendono l'innovazione sostenibile, affidabile e degna della fiducia che i cittadini ripongono nell'istituzione giudiziaria. Preservarli non è un atto di conservatorismo: è un atto di responsabilità verso il futuro della giustizia.

16. Istruzioni operative per l'uso di Copilot e sistemi analoghi: tecniche di prompting

L'efficacia di un sistema di intelligenza artificiale generativa – e la sua conformità alle regole illustrate nelle sezioni precedenti – dipende in misura determinante dalla qualità delle istruzioni che l'operatore gli impartisce. Il prompting non è un'operazione tecnica riservata agli specialisti: è una competenza comunicativa che si apprende, si affina e si governa con la stessa attenzione critica che si dedica alla formulazione di un quesito giuridico. La presente sezione illustra, in forma discorsiva e con suggerimenti tecnici specifici per Microsoft Copilot, le modalità operative raccomandate per ciascuna delle fasi di lavoro tipiche dell'ufficio giudiziario.

L'approccio delineato trova il proprio fondamento valoriale nella Raccomandazione UNESCO sull'etica dell'intelligenza artificiale del 23 novembre 2021 (Res. 41 C/17), che pone la supervisione umana, la trasparenza e la responsabilità individuale come condizioni irrinunciabili di ogni utilizzo istituzionale di questi strumenti.

16.1. Il meta-prompt di sistema: configurare il modello prima di usarlo

Il primo accorgimento tecnico – e il più efficace tra quelli a disposizione dell'operatore – consiste nell'aprire ogni nuova sessione di lavoro con un'istruzione di configurazione che fissi in anticipo le regole del gioco. In Copilot questo si realizza incollando, prima di qualsiasi altro testo, un blocco di istruzioni preliminari che i tecnici del settore chiamano meta-prompt o system prompt: un testo che non produce un output direttamente visibile, ma che orienta il comportamento del modello per tutta la durata della sessione.

Il meta-prompt deve contenere almeno tre tipi di vincoli. Il primo è il vincolo anti-allucinazione: occorre istruire esplicitamente il sistema a non fornire mai numeri di sentenza, date o virgolettati che non siano letteralmente leggibili nelle fonti reperite durante la sessione. Se il dato non è verificabile, il sistema deve dichiararlo con una formula standardizzata, ad esempio apponendo la dicitura «[numero da verificare su banca dati ufficiale]» o «[principio ricostruito – testo da verificare su fonte primaria]». Il secondo è il vincolo di ruolo: il sistema deve essere

qualificato come ausiliario strettamente subordinato al controllo umano, non come sostituto del ragionamento del magistrato. Il terzo è il vincolo di stile: sobrietà terminologica, precisione argomentativa, formato strutturato con una sezione finale dedicata ai riscontri delle fonti citate.

Suggerimento tecnico per Copilot. In Microsoft 365 Copilot è possibile salvare il meta-prompt come elemento riutilizzabile nella sezione “Prompt” del pannello laterale, oppure come primo messaggio fisso di una chat dedicata. In alternativa, si può creare in Word un documento di testo con il meta-prompt già redatto, da incollare all’inizio di ogni nuova sessione con un singolo Ctrl+V. Questa operazione richiede meno di dieci secondi e riduce significativamente la probabilità di risposte inaffidabili.

Un accorgimento ulteriore: quando si apre una nuova scheda di Copilot, il modello non ha memoria delle istruzioni impartite nella sessione precedente. Occorre pertanto premettere il meta-prompt ogni volta, senza dare per scontato che il sistema ricordi le regole già concordate. La dimenticanza di questo passaggio è la causa più frequente di output inaffidabili nella pratica quotidiana.

16.2. Primo esame degli atti: organizzare senza valutare

Quando si utilizza Copilot per la lettura iniziale di un fascicolo, il principio guida è uno solo: il sistema organizza, l’operatore valuta. Nessuna istruzione deve chiedere al modello di identificare ciò che è «rilevante», «importante» o «decisivo» ai fini del giudizio: queste sono operazioni valutative che appartengono esclusivamente alla sfera del ragionamento del magistrato.

Per ottenere una sintesi strutturata degli atti, l’istruzione efficace è quella che delimita con precisione il perimetro dell’operazione. Un prompt ben costruito per questa fase deve: indicare il tipo di documento fornito; specificare che il testo è stato previamente anonimizzato; richiedere un output articolato in sezioni predeterminate (ad esempio: parti e domande, riferimenti normativi espressamente citati, conclusioni rassegnate, elenco dei motivi come rubricati dall’autore); vietare espressamente qualsiasi aggiunta di fatti non presenti nel testo e qualsiasi inferenza sulla fondatezza delle posizioni.

Formulazione raccomandata per la sintesi estrattiva: «Analizza il documento che segue, previamente anonimizzato. Estrai e riporta in forma di elenco strutturato: a) le parti e le domande come indicate nel testo; b) i riferimenti normativi espressamente menzionati dall’autore; c) le conclusioni rassegnate nel dispositivo; d) i motivi di impugnazione o di doglianza con la numerazione e la rubrica originali. Non aggiungere fatti non presenti, non

esprimere valutazioni sulla fondatezza delle posizioni, non selezionare ciò che ritieni rilevante.»

Per l'estrazione di massime giurisprudenziali da provvedimenti ostensibili, l'istruzione deve assegnare al sistema il ruolo del massimatore – non del commentatore – e imporgli i vincoli formali classici: periodo unico sintatticamente unitario, linguaggio impersonale in terza persona, assenza di riferimenti alle parti e di richiami a giurisprudenza esterna. Se il provvedimento non contiene principi generalizzabili, il sistema deve dichiararlo esplicitamente anziché forzare una massimazione artificiale.

Suggerimento tecnico per Copilot. Quando si carica un documento Word o PDF direttamente nella chat di Copilot, il modello elabora il testo del file allegato. Per i fascicoli di grandi dimensioni, è preferibile suddividere l'input per atti distinti anziché caricare un unico documento voluminoso: la qualità dell'output degrada all'aumentare della lunghezza del contesto, soprattutto per le parti finali del documento. La funzione "Riepiloga questo documento" di Copilot in Word è una sintesi astrattiva automatica: non è conforme ai criteri di questa guida e non deve essere utilizzata in sostituzione di un prompt estrattivo esplicito.

16.3. Ricerca giurisprudenziale e studio del diritto: mappare senza selezionare

La ricerca giurisprudenziale assistita da intelligenza artificiale è ammessa, nella fase transitoria, come ausilio per la costruzione di query, la mappatura degli orientamenti e la verifica di precedenti già noti all'operatore. Non è ammessa come strumento di selezione autonoma della giurisprudenza rilevante ai fini della decisione: tale funzione rimane integralmente riservata al magistrato.

Per questa fase è indispensabile premettere il meta-prompt descritto al 1. Senza tale configurazione iniziale, il rischio di citazioni apocrife è elevato e la fiducia nell'output non è giustificata. Va ricordato con fermezza che Copilot, come tutti i modelli generativi, può formulare numeri di sentenza del tutto inesistenti con la stessa fluente sicurezza con cui riporta quelli reali: la plausibilità formale dell'output non è un indice di correttezza sostanziale.

Formulazione raccomandata per la mappatura: «Effettua una ricognizione degli orientamenti su [TEMA], con riferimento a Cassazione (anche Sezioni Unite), Corte Costituzionale e giurisprudenza di merito rilevante. Articola la risposta in: orientamento maggioritario, orientamento minoritario, obiter dicta significativi. Riporta i principi con estremi completi; usa le virgolette solo per testi letteralmente reperiti. Affianca una rassegna sintetica della dottrina accreditata. Concludi con un quadro sinottico delle frizioni ermeneutiche, senza esprimere

valutazioni di merito. Prima di rispondere, esegui un self-audit sulla correttezza degli estremi indicati.»

Per la verifica di un singolo precedente già noto, l'istruzione deve essere costruita in modo da valorizzare la capacità del sistema di accedere al web in tempo reale. Copilot con connessione internet può ricercare attivamente la pronuncia su DeJure, Italgiure o sui siti ufficiali delle corti; occorre però richiedere esplicitamente che, in caso di esito negativo, il sistema dichiari «pronuncia non reperita» senza inventare dati alternativi. La formula di chiusura «esito di verifica puramente fattuale» impedisce al modello di aggiungere valutazioni personali sul peso del precedente.

Per le ricerche su Italgiure, la costruzione della stringa di ricerca richiede un approccio specifico che trae vantaggio dalle caratteristiche del motore: operatore AND implicito, divieto di virgolettare preposizioni articolate elise, massimo sei parole totali, virgolette riservate alle sole locuzioni inscindibili. Chiedere a Copilot di «convertire» un quesito in linguaggio naturale in una stringa Italgiure-compatibile è un uso legittimo e tecnicamente efficace, a condizione che la stringa risultante sia verificata dall'operatore prima dell'immissione.

Suggerimento tecnico per Copilot. Nella versione Microsoft 365 Copilot con accesso a Bing, è possibile richiedere esplicitamente una ricerca web con la formula 'Cerca sul web...' all'inizio del prompt. Senza questa indicazione, il modello risponde prevalentemente attingendo alla propria base di addestramento, che ha una data di aggiornamento non necessariamente recente. Per la giurisprudenza di legittimità degli ultimi due anni è pertanto essenziale indicare esplicitamente 'Cerca sul web le pronunce della Cassazione del 2024 e del 2025 su [tema]' per ottenere risultati aggiornati.

16.4. Supporto alla redazione: strutturare senza decidere

L'utilizzo di Copilot in fase redazionale è forse quello più delicato, perché è quello in cui il confine tra ausilio consentito e delega vietata è più sottile e più facile da valicare per inerzia. La regola di orientamento è semplice: il sistema può elaborare strutture, non può scegliere esiti. Può proporre percorsi argomentativi alternativi equipollenti, non può indicare quale percorso seguire.

Per la bozza strutturale di un provvedimento, l'istruzione deve richiedere espressamente l'articolazione in sezioni distinte – esposizione dei fatti, motivi in diritto, ipotesi motivazionali alternative – e imporre che le ipotesi alternative (accoglimento e rigetto) siano formulate con identica solidità argomentativa. La formula «in modo tecnicamente equipollente e senza

privilegiarne alcuna», inserita nel prompt, non è una cautela formale: è il presidio che impedisce al sistema di orientare, anche implicitamente, la decisione del magistrato. La scelta tra le ipotesi spetta esclusivamente a chi esercita la funzione giurisdizionale.

Formulazione raccomandata per la bozza strutturale: «Elabora una bozza strutturale di [TIPO PROVVEDIMENTO] sul tema [DESCRIZIONE ASTRATTA]. Articola la bozza in tre sezioni: concisa esposizione dei fatti, motivi in punto di diritto, ipotesi motivazionali alternative. Sviluppa un iter argomentativo coerente fondato sulle fonti che indico. Formula le ipotesi alternative (accoglimento e rigetto) in modo tecnicamente equipollente, senza privilegiarne alcuna. La scelta tra le ipotesi è di esclusiva competenza del magistrato. Non inserire dati di fatto che non ti abbia fornito. Riferimenti normativi e giurisprudenziali da sviluppare: [ELENCO].»

Per la revisione stilistico-formale di una bozza già redatta, Copilot è particolarmente efficace se l'istruzione specifica con precisione il tipo di intervento richiesto. I quattro assi di intervento raccomandati sono: l'emendazione delle espressioni colloquiali a favore di un registro formale-istituzionale; la ricomposizione della consequenzialità logico-sintattica dei periodi eccessivamente contorti; la verifica e la segnalazione dei richiami giurisprudenziali generici o non localizzati; la redazione di una nota redazionale finale che elenchi gli interventi effettuati. Quest'ultimo elemento è di particolare importanza ai fini della trasparenza: consente al magistrato di valutare con esattezza in che misura il testo è stato modificato rispetto all'originale.

Un vincolo che non deve mai mancare nel prompt di revisione è il divieto di modificare il contenuto sostanziale e le scelte argomentative dell'autore. Senza questa esplicita limitazione, i modelli generativi tendono spontaneamente a «reiscrivere» anziché a «revisare», alterando la struttura logica e talvolta introducendo argomenti non presenti nell'originale. La distinzione tra revisione e riscrittura non è banale: deve essere imposta esplicitamente, ogni volta.

Suggerimento tecnico per Copilot in Word. La funzione 'Riscrivi con Copilot' disponibile nel menú contestuale di Word opera una riscrittura automatica e non vincolata del testo selezionato: non è conforme ai criteri di questa guida e non deve essere usata su bozze di provvedimenti. Per una revisione controllata, è preferibile aprire il pannello laterale di Copilot, incollare manualmente il testo da revisionare e formulare l'istruzione con i vincoli espliciti descritti in questa sezione. In questo modo l'operatore mantiene il controllo sull'intero processo.

16.5. Regole di condotta generale nel prompting

Al di là delle istruzioni specifiche per ciascuna fase operativa, vi sono alcune regole di condotta generale che presiedono a ogni interazione con Copilot o con sistemi analoghi e che derivano

direttamente dai principi enunciati nelle sezioni precedenti del presente vademecum.

La prima è la regola della specificità: un prompt vago produce un output vago, e un output vago non è verificabile. Ogni istruzione deve indicare con precisione il ruolo attribuito al sistema, l'obiettivo dell'operazione, i vincoli di contenuto e di forma, il formato atteso della risposta. La specificità non è un appesantimento del lavoro: è la condizione che rende il risultato utilizzabile.

La seconda è la regola del controllo iterativo: l'output di una singola interrogazione non è mai definitivo. Il dialogo con il sistema deve essere concepito come un processo a più passaggi – si verifica, si corregge, si chiede di approfondire, si richiede la fonte di un'affermazione specifica. La formula «mostrami la fonte di questa affermazione», aggiunta come follow-up, è uno degli strumenti di controllo più efficaci a disposizione dell'operatore: costringe il sistema a esplicitare il proprio riferimento o a dichiarare di non averne uno.

La terza è la regola della separazione degli input: non si mescola in un unico prompt la funzione di ricerca e la funzione di redazione. Se si chiede al sistema di «cercare la giurisprudenza e poi scrivere la motivazione», il modello tende a produrre una risposta in cui i confini tra ciò che ha trovato e ciò che ha elaborato autonomamente sono indistinguibili. Le due operazioni devono essere condotte in sessioni o in messaggi separati, con verifica intermedia dell'output della ricerca prima di procedere alla fase redazionale.

La quarta è la regola della tracciabilità: ogni utilizzo operativamente rilevante deve poter essere ricostruito. Ai fini della documentazione prescritta nella Sezione X, è sufficiente conservare: il testo del prompt principale utilizzato, l'output ricevuto, la nota di verifica redatta dall'operatore. In Copilot la cronologia delle conversazioni è accessibile dal pannello laterale: è buona prassi attribuire a ogni sessione un titolo descrittivo che ne consenta il reperimento successivo.

La quinta è la regola della riservatezza attiva: non si aspetta di commettere un errore per ricordare che i dati riservati non devono essere inseriti. Prima di incollare qualsiasi testo nel prompt, l'operatore deve verificare attivamente che non contenga dati identificativi, numeri di procedimento, nomi di parti o di persone fisiche, atti coperti da segreto. La verifica preliminare è un atto consapevole, non un controllo residuale.

Fondamento UNESCO. La Raccomandazione sull'etica dell'intelligenza artificiale (UNESCO, 23 novembre 2021, Res. 41 C/17) afferma che la supervisione umana significativa – intesa non come mera presenza formale ma come controllo effettivo, informato e critico sull'output dei sistemi – è

la condizione che distingue un utilizzo etico dell'IA da una delega occulta della funzione. Le regole operative descritte in questa sezione non sono un insieme di precauzioni burocratiche: sono la traduzione pratica di quel principio nel contesto specifico della funzione giurisdizionale.

Riferimenti bibliografici

Consiglio Superiore della Magistratura, ***Raccomandazioni sull'uso dell'intelligenza artificiale nell'amministrazione della giustizia***, delibera plenaria 8 ottobre 2025.

F. Iannone, ***Prime riflessioni sull'IA: la legge 1232/2025 letta in cornice eurounitaria e sovranazionale***, in Giustizia Insieme, 7 novembre 2025.

Testo integrale e navigazione interattiva del **Regolamento (UE) 2024/1689 (AI Act)**.

AGID – Agenzia per l'Italia Digitale e **ACN – Agenzia per la Cybersicurezza Nazionale**, autorità nazionali designate ai sensi dell'art. 20 del Regolamento (UE) 2024/1689.

Regolamento (UE) n. 910/2014 del Parlamento europeo e del Consiglio (eIDAS), art. 46, in G.U. UE L 257 del 28 agosto 2014.

CEN – Comitato Europeo di Normazione e CENELEC – Comitato Europeo di Normazione Elettrotecnica, organismi europei di normalizzazione riconosciuti dall'Unione europea per lo sviluppo degli standard armonizzati ai sensi del Regolamento (UE) 2024/1689.

F. Iannone, *IA e giurisdizione: le prime riflessioni*, in Gazzetta Forense, n. 4/2025.