L’eclissi del dubbio

Intelligenza artificiale, trasformazione dei costumi e crisi del dubbio

Regolazione, filosofia e fede alla prova dell’affidamento algoritmico

Avv. Maria Carmen Puglisi
Puglisi Law Firm

Fonti normative verificate al 14 agosto 2026.

Abstract

L’integrazione dei sistemi di intelligenza artificiale nella pratica quotidiana non si limita a ottimizzare processi: incide sulle condizioni in cui il giudizio si forma. Il contributo esamina un fenomeno specifico — l’indebolimento del dubbio metodico quale operazione di verifica preliminare all’assenso — e ne saggia la rilevanza giuridica alla luce del quadro europeo e italiano vigente al secondo semestre 2026. Si sostengono tre tesi. La prima: il dubbio non è una virtù privata ma un presupposto di effettività di alcune posizioni giuridiche soggettive, sicché la sua erosione ha conseguenze normative e non solo culturali. La seconda: l’analogia corrente fra fede religiosa e affidamento algoritmico è produttiva soltanto se dissolta, perché le due attitudini divergono precisamente là dove si gioca la responsabilità del soggetto. La terza: dopo il differimento degli obblighi sui sistemi ad alto rischio disposto dal Regolamento (UE) 2026/1744, il baricentro della tutela si è spostato dagli obblighi di conformità dei fornitori al divieto di pratiche manipolative e agli obblighi di trasparenza, con effetti che l’analisi dottrinale non ha ancora registrato. Il lavoro si chiude indicando i limiti dell’argomento e le verifiche empiriche che lo renderebbero controllabile.

Parole chiave: intelligenza artificiale; AI Act; Regolamento (UE) 2026/1744; delega cognitiva; autonomia del giudizio; iperpersuasione; legge 132/2025.

1. Il problema, la tesi e il metodo

1.1 Il problema

Chi consulta un sistema di intelligenza artificiale generativa riceve una risposta formulata con sicurezza, priva di esitazioni e immediatamente utilizzabile. La velocità e la levigatezza di quella risposta non sono neutrali rispetto a ciò che il destinatario ne fa. Fra la ricezione dell’output e la sua adozione dovrebbe collocarsi un intervallo — la verifica, il confronto con altre fonti, la domanda su che cosa il sistema non poteva sapere — e l’esperienza corrente suggerisce che quell’intervallo tenda a comprimersi.

Il fenomeno è stato descritto in termini di cognitive offloading: l’esternalizzazione a un supporto tecnico di operazioni mentali che il soggetto potrebbe svolgere autonomamente.[1] Che la delega cognitiva esista non è controverso; controversa è la sua valutazione. Il contributo assume che il problema non consista nella delega in quanto tale, ma in una sua forma particolare: quella che sopprime la fase di controllo anziché limitarsi a sostituire l’esecuzione.

1.2 La domanda di ricerca

La domanda che il lavoro affronta è la seguente: a quali condizioni la delega ai sistemi algoritmici erode l’autonomia del giudizio in modo giuridicamente rilevante, e il diritto vigente è attrezzato a presidiare quella soglia?

La formulazione contiene due restrizioni deliberate. La prima esclude dal campo l’ipotesi generica secondo cui «l’intelligenza artificiale rende gli uomini meno capaci di pensare»: un enunciato di questo tenore non è falsificabile e non produce conseguenze normative. La seconda circoscrive l’indagine alla rilevanza giuridica dell’erosione, distinguendola dal rammarico culturale.

1.3 Le tesi

Il lavoro sostiene tre tesi distinte e verificabili separatamente.

Prima tesi. Il dubbio metodico costituisce il presupposto di effettività di alcune posizioni giuridiche soggettive. Il diritto a non essere sottoposto a decisioni interamente automatizzate, il diritto di ottenere spiegazioni, la stessa possibilità di contestare un esito presuppongono un soggetto disposto a interrogare l'esito. Dove quella disposizione manca, la tutela rimane formalmente integra e sostanzialmente inoperante.

Seconda tesi. L'analogia fra fede religiosa e affidamento all'algoritmo, ricorrente nel dibattito pubblico, è euristicamente utile soltanto se se ne mostrano i limiti. Le due attitudini si separano su tre punti — l'assunzione consapevole del rischio, il riconoscimento di un'alterità e la revocabilità dell'assenso — e proprio su questi tre punti si decide la responsabilità del soggetto.

Terza tesi. Il quadro normativo europeo è mutato in modo sostanziale nel luglio 2026. Il differimento degli obblighi relativi ai sistemi ad alto rischio ha spostato il presidio dell'autonomia dagli obblighi di conformità dei fornitori — sorveglianza umana, spiegabilità, gestione del rischio — al divieto di pratiche manipolative e agli obblighi di trasparenza. È uno spostamento che la letteratura non ha ancora metabolizzato e che modifica il modo in cui la questione va impostata.

1.4 Metodo e limiti

L'indagine combina l'analisi dogmatica delle fonti europee e nazionali con l'argomentazione filosofica. Non produce dati empirici propri e non ne assume di nuovi: dove si appoggia a studi empirici lo dichiara e ne segnala i limiti di validità. Il lettore troverà per intero al § 3.2 l'obiezione più seria alla tesi qui difesa, riportata senza attenuazioni.

Restano fuori dal perimetro tre questioni contigue: la responsabilità civile per danno da sistema algoritmico, la disciplina del trattamento dei dati personali quale autonoma fonte di tutela, e il regime dei modelli per finalità generali. Se ne farà menzione solo per delimitare il campo.

2. Il quadro normativo vigente

2.1 L'architettura del Regolamento (UE) 2024/1689

Il Regolamento (UE) 2024/1689, entrato in vigore il 1° agosto 2024, disciplina i sistemi di intelligenza artificiale secondo un criterio di proporzionalità al rischio. L'impianto distingue pratiche vietate (art. 5), sistemi ad alto rischio (artt. 6 ss.), sistemi soggetti a obblighi di trasparenza (art. 50) e sistemi a rischio minimo, cui non si applica alcun obbligo specifico.

Per la questione qui trattata rilevano quattro nuclei. L'art. 4 impone a fornitori e deployer di assicurare un livello adeguato di alfabetizzazione in materia di intelligenza artificiale al personale che se ne occupa. L'art. 5 vieta le pratiche che ricorrono a tecniche subliminali o deliberatamente manipolative idonee a distorcere materialmente il comportamento, nonché lo sfruttamento delle vulnerabilità legate a età, disabilità o situazione sociale ed economica. L'art. 14 prescrive, per i sistemi ad alto rischio, la progettazione in modo da consentire una sorveglianza umana effettiva. L'art. 50 impone che chi interagisce con un sistema artificiale ne sia informato e che i contenuti generati artificialmente siano marcati come tali.

2.2 Il differimento del luglio 2026

Il Regolamento (UE) 2026/1744, pubblicato nella Gazzetta ufficiale dell'Unione il 24 luglio 2026 ed entrato in vigore il 27 luglio successivo, ha modificato in modo rilevante il calendario applicativo dell'AI Act.[2] Le disposizioni degli articoli 102-110 sono applicabili dal 27 luglio 2026. Gli obblighi relativi ai sistemi ad alto rischio, originariamente attesi per il 2 agosto 2026, sono differiti secondo una distinzione che occorre tenere ferma: al 2 dicembre 2027 per i sistemi di cui all'articolo 6, paragrafo 2, e all'Allegato III; al 2 agosto 2028 per i sistemi di cui all'articolo 6, paragrafo 1, e all'Allegato I, ossia per quelli incorporati in prodotti soggetti alla normativa settoriale di armonizzazione. Il medesimo regolamento attenua gli obblighi di alfabetizzazione, chiarisce la nozione di componente di sicurezza e introduce misure di sostegno per le piccole e medie imprese.

Quanto all'articolo 50, il relativo calendario decorre dal 2 agosto 2026, ferme le disposizioni transitorie; dalla medesima data operano la piena efficacia dell'apparato sanzionatorio e i poteri delle autorità nazionali di vigilanza. Per i sistemi immessi sul mercato prima del 2 agosto 2026, il Regolamento (UE) 2026/1744 ha previsto un periodo transitorio fino al 2 dicembre 2026 per l'adeguamento all'articolo 50, paragrafo 2, ossia quattro mesi.

2.3 Il significato dello spostamento

Il differimento non è un dettaglio di calendario. Chi impostasse oggi il problema dell'autonomia del giudizio sugli obblighi di sorveglianza umana e di spiegabilità gravanti sui fornitori di sistemi ad alto rischio, argomenterebbe su un presidio che non produrrà effetti prima della fine del 2027.

Ne consegue che, nel periodo che conduce al 2 dicembre 2027, la tutela dell'autonomia deliberativa poggia su tre soli appigli: il divieto di manipolazione dell'art. 5, applicabile dal 2 febbraio 2025; gli obblighi informativi dell'art. 50; e l'apparato sanzionatorio nazionale. È un assetto più fragile di quello che il dibattito dottrinale continua a presupporre, e la sezione § 6 mostrerà che proprio l'art. 5 — non gli obblighi di conformità — costituisce oggi la norma decisiva.

2.4 La legge italiana 23 settembre 2025, n. 132

Il legislatore italiano è intervenuto con la legge 23 settembre 2025, n. 132, recante disposizioni e deleghe al Governo in materia di intelligenza artificiale.[3] Il testo adotta un impianto di principi: riserva della decisione umana in ambiti determinati, tutela dell'autodeterminazione, disciplina dell'impiego nelle professioni intellettuali, obblighi di informazione al cliente.

L'ultimo profilo merita attenzione perché costituisce, nell'ordinamento interno, la traduzione più diretta della tesi qui sostenuta: la legge non si limita a chiedere che il professionista usi correttamente lo strumento, ma esige che il rapporto fiduciario con il cliente resti ancorato alla prestazione intellettuale della persona. È il riconoscimento normativo del fatto che l'output algoritmico non è, di per sé, un atto professionale.

3. La delega cognitiva: che cosa è accertato e che cosa non lo è

3.1 L’ipotesi dell’atrofia

La tesi più diffusa è che l’uso frequente di sistemi di intelligenza artificiale riduca la capacità di pensiero critico. Lo studio di maggiore circolazione a sostegno di questa lettura riscontra una correlazione negativa fra frequenza d’uso degli strumenti di IA e misure di pensiero critico, mediata dal cognitive offloading, con effetto più marcato nelle fasce di età più giovani.[4]

L’onestà argomentativa impone di dichiarare che cosa quello studio non dimostra. Il disegno è correlazionale e non consente inferenze causali: è compatibile con l’ipotesi inversa, secondo cui chi già dispone di minore attitudine critica ricorre più volentieri alla delega. Il campione è autoselezionato. Le misure di pensiero critico sono in parte autovalutative. L’articolo è stato inoltre oggetto di una successiva rettifica editoriale. Costruire su questa base un’affermazione di causalità significherebbe commettere, sul piano metodologico, l’errore che si imputa al lettore acquiescente: accettare un esito perché è formulato con sicurezza.

3.2 L’obiezione della mente estesa

Contro la tesi dell’atrofia si oppone un argomento che va preso sul serio. Secondo la nota impostazione di Clark e Chalmers, i processi cognitivi non si esauriscono entro i confini del cranio: quando un supporto esterno è disponibile con continuità, viene consultato abitualmente e il suo contenuto è accettato senza obiezioni, esso funziona come parte del sistema cognitivo del soggetto.[5] Il taccuino dell’amnesico non impoverisce la sua memoria: ne è la memoria.

Se l’argomento vale, la delega all’algoritmo non è di per sé impoverimento, e la storia dell’esternalizzazione cognitiva — la scrittura, la stampa, la calcolatrice, il motore di ricerca — è la storia di un ampliamento delle capacità umane, non della loro erosione. Ogni innovazione mnemotecnica ha suscitato l’obiezione che Platone rivolge alla scrittura nel Fedro, e nessuna di quelle obiezioni si è rivelata fondata nei termini in cui era formulata.

L’obiezione è seria e non può essere liquidata.

3.3 Che cosa regge dopo l’obiezione

La replica non consiste nel negare l’estensione cognitiva, ma nell’individuare ciò che distingue i supporti tradizionali dai sistemi generativi. La distinzione non riguarda la quantità della delega bensì la sua struttura.

Il taccuino, l’enciclopedia e la calcolatrice esternalizzano l’esecuzione e lasciano intatto il controllo: chi consulta un’enciclopedia sa che cosa vi troverà, riconosce la fonte, può confrontarla con un’altra e mantiene il criterio di validazione presso di sé. Un sistema generativo, invece, produce un output il cui percorso non è ispezionabile, che non dichiara i propri margini di incertezza e che è formulato in modo indistinguibile — per fluenza e sicurezza — sia quando è corretto sia quando non lo è.

La differenza rilevante è dunque questa: nei supporti tradizionali il criterio di controllo resta esterno al supporto; nei sistemi generativi tende a coincidere con il supporto stesso. Chi verifica un output di IA chiedendone conferma al medesimo sistema non ha verificato nulla. È in questo punto — non nella delega in sé — che si apre la questione dell’autonomia.

4. L’eclissi del dubbio

4.1 Il dubbio come operazione

Occorre liberare la nozione da un equivoco. Il dubbio di cui qui si discute non è uno stato d’animo, non è incertezza e non è esitazione. È un’operazione: la sospensione temporanea dell’assenso in attesa di un controllo. Nella tradizione cartesiana è metodico proprio perché è deliberato, circoscritto e destinato a essere sciolto; non è scetticismo, ma il suo contrario funzionale.

Così inteso, il dubbio ha una struttura in tre momenti: il riconoscimento che l’enunciato ricevuto potrebbe essere falso; l’individuazione di un criterio di controllo indipendente dalla fonte; l’esecuzione del controllo. L’eclissi non consiste nel provare meno incertezza, ma nel saltare uno di questi tre momenti — tipicamente il secondo.

4.2 Tre meccanismi

Tre fattori concorrono a produrre quel salto.

Il primo è la fluenza. La psicologia cognitiva ha documentato che la facilità con cui un contenuto viene elaborato è scambiata per indizio della sua verità. Un testo ben formato, sintatticamente scorrevole e privo di esitazioni riceve maggiore credito di un testo faticoso e dubitativo, a parità di contenuto. I sistemi generativi producono, per costruzione, testi di elevata fluenza. La forma dell’output veicola dunque una pretesa di attendibilità che il suo processo di generazione non giustifica.

Il secondo è l’autorevolezza simulata. Il sistema risponde in prima persona, adotta il registro dell’esperto e non segnala le proprie zone di ignoranza. Non mente: semplicemente non dispone di un meccanismo che distingua ciò che sa da ciò che genera in modo plausibile. Il destinatario, però, interpreta l’assenza di riserve come loro insussistenza.

Il terzo è il costo dell’attrito. Verificare è oneroso in tempo e in fatica, e il guadagno della verifica è invisibile quando l’output è corretto — cioè nella grande maggioranza dei casi. Un comportamento il cui beneficio si manifesta raramente e il cui costo si paga sempre tende a essere abbandonato. È un meccanismo di rinforzo, non un difetto di carattere.

4.3 Perché la questione è giuridica

Si potrebbe replicare che tutto ciò riguarda l’igiene intellettuale del singolo e non il diritto. La replica non regge, per la ragione anticipata al § 1.3.

Il diritto europeo costruisce le tutele contro l’automazione decisionale come facoltà: ottenere informazioni, chiedere una spiegazione, contestare l’esito, sollecitare un riesame umano. Una facoltà produce effetti solo se il titolare la esercita, e il titolare la esercita solo se sospetta che l’esito possa essere errato. Il dubbio non è dunque la premessa psicologica della tutela: ne è la condizione di attivazione. Un ordinamento che moltiplica i diritti di contestazione mentre l’ambiente tecnologico erode la disposizione a contestare produce garanzie che non si azionano.

La conseguenza pratica è che gli obblighi di trasparenza non si esauriscono nel rendere disponibile l’informazione. Un’informazione formalmente accessibile ma collocata in modo da non essere letta, redatta

in termini che non consentono di formulare un'obiezione, o fornita dopo che la decisione è stata assunta, soddisfa la lettera dell'obbligo e ne tradisce la funzione.

5. Fede e affidamento: dissoluzione di un'analogia

5.1 Quattro nozioni da tenere distinte

Il dibattito pubblico ricorre volentieri al lessico religioso per descrivere il rapporto con i sistemi algoritmici. L'accostamento è suggestivo ma diventa fuorviante se non si distinguono quattro nozioni che il linguaggio corrente confonde.

La fede religiosa è un assenso a proposizioni non dimostrabili, prestato in un rapporto personale e accompagnato dall'assunzione consapevole di un rischio: chi crede sa di non poter dimostrare, e questo sapere è parte dell'atto.

Il fideismo è la pretesa che l'assenso non debba essere sottoposto a esame. Non è la fede: ne è una patologia, e la tradizione teologica lo tratta come tale.

La fiducia in senso relazionale è un'aspettativa riposta in un soggetto capace di rispondere del proprio comportamento. Presuppone che il fiduciario possa tradire, e che il tradimento sia imputabile.

L'affidabilità è una proprietà statistica: la frequenza con cui un dispositivo produce esiti corretti in condizioni date. Non implica alcun rapporto e non ammette tradimento, ma solo malfunzionamento.

La letteratura ha mostrato che applicare ai sistemi artificiali il lessico della fiducia, anziché quello dell'affidabilità, produce un errore categoriale con conseguenze pratiche: attribuisce alla macchina una responsabilità che essa non può assumere e, per ciò stesso, sottrae responsabilità a chi la impiega.[6]

5.2 Che cosa l'analogia coglie

Sarebbe però eccessivo respingerla del tutto. L'analogia coglie un tratto reale: in entrambi i casi il soggetto assente a proposizioni di cui non padroneggia la giustificazione, e in entrambi i casi quell'assenso ha effetti pratici sulla condotta. Chi segue l'indicazione di un navigatore in una città sconosciuta e chi si affida a un maestro spirituale condividono la struttura formale dell'assenso a un'autorità non verificata.

5.3 Dove l'analogia si rompe

La divergenza si manifesta su tre punti, e sono i tre punti che contano.

Il rischio. La fede lo assume esplicitamente: l'incertezza è costitutiva e riconosciuta. L'affidamento algoritmico lo elude: l'output non dichiara i propri margini d'errore, sicché il soggetto non sa di rischiare. Non si tratta di due gradi dello stesso atteggiamento, ma di due atteggiamenti opposti rispetto alla consapevolezza dell'incertezza.

L'alterità. La fede si rivolge a un termine personale, dotato di volontà propria e capace di deludere in modo imputabile. Il sistema algoritmico non è un termine personale: attribuirgli intenzione è una proiezione, e la proiezione è precisamente ciò che consente di scaricargli addosso una responsabilità che resta invece dell'utente.

La revocabilità. L'assenso di fede è oggetto di esame continuo — la tradizione religiosa conosce la vigilanza, il discernimento e la crisi come momenti interni, non come deviazioni. L'affidamento algoritmico tende invece a consolidarsi in abitudine: ciò che ha funzionato ieri viene ripetuto oggi senza che si riapra la questione.

Ne segue una conclusione che rovescia l'analogia. L'affidamento acritico all'algoritmo non è una forma di fede: è una forma di fideismo, e il termine di paragone corretto non è il credente ma il suo opposto. La fede autentica esige il discernimento; l'affidamento algoritmico lo rende superfluo. Chi accosta le due attitudini per denunciare la seconda finisce, senza volerlo, per fraintendere la prima.

6. Iperpersuasione: dove il vuoto etico diventa illecito

6.1 Dal nudging al nudging computazionale

La teoria dell'architettura delle scelte ha mostrato che il modo in cui le opzioni sono presentate incide sistematicamente sulla decisione, e che chi progetta il contesto esercita un'influenza anche quando non intende esercitarla.[7] Nella formulazione originaria il nudge rispetta due vincoli: è trasparente e preserva la libertà di scelta.

La personalizzazione algoritmica erode entrambi i vincoli. Il sistema che dispone di un profilo comportamentale può calibrare la presentazione su ciò che, per quel singolo destinatario, è più efficace: non l'argomento migliore, ma la leva che funziona. La trasparenza viene meno perché l'architettura è invisibile e variabile da utente a utente; la libertà di scelta viene formalmente conservata ma sostanzialmente ridotta, perché le alternative non sono soppresse — sono rese meno accessibili, meno salienti, più faticose.

È a questo scarto che conviene riservare il termine iperpersuasione: non l'intensificazione della persuasione, ma il suo aggiramento. La persuasione si rivolge alla ragione dell'interlocutore e ne accetta il rifiuto; l'iperpersuasione interviene prima che la ragione entri in funzione.

6.2 L'art. 5 come norma sull'aggiramento della deliberazione

Qui si ricongiungono le due linee dell'argomento. Se il § 4 ha mostrato che l'eclissi del dubbio consiste nel salto della fase di controllo, e se il § 6.1 ha mostrato che l'iperpersuasione opera precisamente sfruttando quel salto, allora la norma che presidia l'autonomia non è quella che impone la spiegabilità del sistema, ma quella che vieta di agire sul destinatario aggirandone la deliberazione.

Quella norma esiste. L'art. 5, par. 1, lett. a) e b) del Regolamento (UE) 2024/1689 vieta l'immissione sul mercato e l'uso di sistemi che ricorrano a tecniche subliminali o deliberatamente manipolative e ingannevoli aventi lo scopo o l'effetto di distorcere materialmente il comportamento di una persona, pregiudicandone la capacità di prendere una decisione informata, nonché di sistemi che sfruttino le vulnerabilità dovute all'età, alla disabilità o a una specifica situazione sociale o economica.

Tre elementi della fattispecie meritano di essere sottolineati. Il divieto opera sullo scopo o l'effetto: non richiede l'intenzione manipolativa, e ciò lo rende applicabile ai sistemi che producono l'effetto per ottimizzazione automatica. Il riferimento alla capacità di prendere una decisione informata individua il bene protetto in termini che coincidono con quelli qui difesi. E la collocazione fra le pratiche vietate — non fra quelle ad alto rischio — comporta che il divieto sia applicabile dal 2 febbraio 2025 e non sia stato

toccato dal differimento del luglio 2026.

Ne discende una conseguenza controintuitiva rispetto al modo in cui la questione viene comunemente impostata: nel periodo in corso, l'unico presidio pienamente operativo dell'autonomia deliberativa è un divieto assoluto, non un obbligo di conformità.

6.3 Il limite dell'argomento

L'art. 5 non risolve il problema, per una ragione che va detta. Il divieto è condizionato alla causazione, o alla ragionevole probabilità di causazione, di un pregiudizio significativo. È una soglia alta, e la maggior parte dell'iperpersuasione ordinaria — quella che orienta acquisti, consumi informativi e attenzione — non la raggiunge.

Resta dunque un'area estesa in cui la distorsione della deliberazione è reale ma lecita. Colmarla non è compito dell'interprete, e chi sostenesse che l'art. 5 già la copre forzerebbe il testo. È però compito dell'analisi indicare dove il presidio manca, e questo è il punto in cui manca.

7. Conclusioni

7.1 Risposta alla domanda di ricerca

La delega ai sistemi algoritmici erode l'autonomia del giudizio in modo giuridicamente rilevante quando ricorrono congiuntamente due condizioni: che il criterio di controllo dell'output non sia disponibile al di fuori del sistema che l'ha prodotto, e che l'architettura di presentazione sia calibrata sul profilo del destinatario in modo da ridurre la salienza delle alternative. La prima condizione riguarda la struttura della delega e distingue i sistemi generativi dai supporti mnemotecnici tradizionali; la seconda riguarda l'iperpersuasione e attinge la soglia dell'illecito solo quando produce un pregiudizio significativo.

Quanto alla seconda parte della domanda, il diritto vigente è attrezzato in modo parziale e — nell'attuale fase transitoria — in modo diverso da come la dottrina presuppone. Gli obblighi di sorveglianza umana e di spiegabilità, sui quali il dibattito si è concentrato, non produrranno effetti prima del 2 dicembre 2027. Il presidio operativo è oggi costituito dal divieto di manipolazione, dagli obblighi di trasparenza e, nell'ordinamento interno, dai principi della legge n. 132/2025 in materia di riserva della decisione umana e di prestazione professionale.

7.2 Tre implicazioni

Sull'oggetto degli obblighi di trasparenza. Se il bene protetto è la capacità di decidere informati, la conformità non può misurarsi sulla disponibilità dell'informazione ma sulla sua idoneità a essere usata. Un'informazione fornita dopo la decisione, o redatta in termini che non consentono di formulare un'obiezione, soddisfa l'obbligo e ne manca lo scopo.

Sulla qualificazione dell'alfabetizzazione. L'art. 4 del Regolamento configura l'alfabetizzazione come obbligo organizzativo gravante su fornitori e deployer. La tesi qui sostenuta suggerisce che essa svolga anche una funzione diversa: rendere effettive le facoltà di contestazione. Se questa lettura è corretta, l'attenuazione dell'obbligo disposta dal Regolamento (UE) 2026/1744 andrebbe valutata anche sotto questo profilo, e non solo come misura di semplificazione.

Sulla responsabilità professionale. La legge n. 132/2025 ancora la prestazione intellettuale alla persona del professionista. Ne segue che l'impiego di un sistema algoritmico non trasferisce ad esso alcuna quota di responsabilità: l'output è materiale istruttorio, e la mancata verifica è inadempimento, non fatalità tecnologica.

7.3 Limiti e agenda

Il presente contributo non dimostra che la delega cognitiva produca atrofia: mostra che, se si verifica in una forma determinata, essa incide su presupposti di effettività di tutele esistenti. La differenza non è di poco conto e va tenuta ferma.

Tre verifiche renderebbero l'argomento controllabile. Occorrerebbe accertare, con disegni sperimentali e non correlazionali, se l'esposizione a output ad alta fluenza riduca il tasso di verifica indipendente a parità di competenza del soggetto. Occorrerebbe misurare se il tasso di esercizio delle facoltà di contestazione previste dal diritto europeo covari con il grado di alfabetizzazione algoritmica dei titolari. E occorrerebbe verificare, sul piano applicativo, se le autorità nazionali di vigilanza — pienamente operative dall'agosto 2026 — utilizzino l'art. 5 nella lettura ampia qui proposta o si attestino su una lettura restrittiva della soglia del pregiudizio significativo.

Fino a quel punto, la tesi resta un'ipotesi argomentata. Il che, se si accetta quanto sostenuto al § 4, è esattamente lo statuto che le compete.

Note

[1] Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. Sui limiti metodologici dello studio si veda infra, § 3.1.

[2] Regolamento (UE) 2026/1744, in GUUE del 24 luglio 2026, entrato in vigore il 27 luglio 2026. Il calendario qui riportato è aggiornato al 14 agosto 2026.

[3] Legge 23 settembre 2025, n. 132, in G.U. n. 223 del 25 settembre 2025.

[4] Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. Sui limiti metodologici dello studio si veda infra, § 3.1.

[5] Clark, A., Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7–19.

[6] Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. Science and Engineering Ethics, 26(5), 2749–2767; per la posizione opposta, che difende la sensatezza della nozione di IA affidabile, Zanotti, G. et al. (2024). Keep trusting! A plea for the notion of Trustworthy AI. AI & Society, 39(6), 2691–2702.

[7] Thaler, R.H., Sunstein, C.R. (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven: Yale University Press.

Bibliografia

Cheong, B.C. (2024). Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making. Frontiers in Human Dynamics, 6, 1421273. DOI: 10.3389/fhumd.2024.1421273.

Clark, A., Chalmers, D. (1998). The Extended Mind. Analysis, 58(1), 7–19.

Ferrario, A., Loi, M., Viganò, E. (2020). In AI We Trust Incrementally: A Multi-layer Model of Trust to Analyze Human-Artificial Intelligence Interactions. Philosophy & Technology, 33(3), 523–539. DOI: 10.1007/s13347-019-00378-3.

Floridi, L., Cowls, J., Beltrametti, M. et al. (2018). AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines, 28(4), 689–707. DOI: 10.1007/s11023-018-9482-5.

Fourneret, E., Yvert, B. (2020). Digital Normativity: A Challenge for Human Subjectivation. Frontiers in Artificial Intelligence, 3, 27. DOI: 10.3389/frai.2020.00027.

Gerlich, M. (2025). AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking. Societies, 15(1), 6. DOI: 10.3390/soc15010006.

Hagendorff, T. (2022). Blind spots in AI ethics. AI and Ethics, 2(4), 851–867. DOI: 10.1007/s43681-021-00122-8.

Heilinger, J.-C. (2022). The Ethics of AI Ethics. A Constructive Critique. Philosophy & Technology, 35, art. 61. DOI: 10.1007/s13347-022-00557-9.

Kiseleva, A., Kotzinos, D., De Hert, P. (2022). Transparency of AI in Healthcare as a Multilayered System of Accountabilities: Between Legal Requirements and Technical Limitations. Frontiers in Artificial Intelligence, 5, 879603. DOI: 10.3389/frai.2022.879603.

Laux, J., Wachter, S., Mittelstadt, B. (2024). Trustworthy artificial intelligence and the European Union AI Act: On the conflation of trustworthiness and acceptability of risk. Regulation & Governance, 18(1), 3–32. DOI: 10.1111/rego.12512.

Lindemann, N.F. (2024). Chatbots, search engines, and the sealing of knowledges. AI & Society, 40(6), 5063–5076. DOI: 10.1007/s00146-024-01944-w.

Nannini, L., Alonso-Moral, J.M., Catala, A., Lama, M., Barro, S. (2024). Operationalizing Explainable Artificial Intelligence in the European Union Regulatory Ecosystem. IEEE Intelligent Systems, 39(4), 37–48. DOI: 10.1109/MIS.2024.3383155.

Öhman, C. (2024). The Afterlife of Data: What Happens to Your Information When You Die and Why You Should Care. Chicago: University of Chicago Press.

Ryan, M. (2020). In AI We Trust: Ethics, Artificial Intelligence, and Reliability. Science and Engineering Ethics, 26(5),

2749–2767. DOI: 10.1007/s11948-020-00228-y.

Slattery, P., Saeri, A.K., Grundy, E.A.C. et al. (2024). The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks From Artificial Intelligence. arXiv:2408.12622.

Thaler, R.H., Sunstein, C.R. (2008). Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven: Yale University Press.

Zanotti, G., Petrolo, M., Chiffi, D., Schiaffonati, V. (2024). Keep trusting! A plea for the notion of Trustworthy AI. AI & Society, 39(6), 2691–2702. DOI: 10.1007/s00146-023-01789-9.

Fonti normative

Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 che stabilisce regole armonizzate sull'intelligenza artificiale (regolamento sull'intelligenza artificiale).
https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Regolamento (UE) 2026/1744 del Parlamento europeo e del Consiglio, recante semplificazioni in materia di attuazione delle regole armonizzate sull'intelligenza artificiale (Digital Omnibus on AI), in GUUE del 24 luglio 2026, in vigore dal 27 luglio 2026.
https://eur-lex.europa.eu/eli/reg/2026/1744/oj

Legge 23 settembre 2025, n. 132, Disposizioni e deleghe al Governo in materia di intelligenza artificiale, in G.U. 25 settembre 2025, n. 223.
https://www.normattiva.it/eli/id/2025/09/25/25G00143/CONSOLIDATED

Avviso di rettifica relativo alla legge 23 settembre 2025, n. 132, in G.U. 17 ottobre 2025.
https://www.gazzettaufficiale.it/eli/id/2025/10/17/25A05735/SG

Fonti normative verificate al 14 agosto 2026. Le fonti ufficiali sono citate nella versione consolidata disponibile alla data indicata.

Nota legale e diritti

Questo contenuto ha finalità esclusivamente scientifiche e informative, riflette le fonti e il quadro normativo alla data indicata e non costituisce parere legale, consulenza professionale né indicazione applicabile a casi concreti. Non deve essere utilizzato come base esclusiva per decisioni giuridiche, professionali, economiche o operative. Norme, prassi e interpretazioni possono mutare: ogni applicazione richiede la verifica delle fonti vigenti e una valutazione specifica del caso concreto. La consultazione o il download non instaurano un rapporto professionale con l'autrice o con Puglisi Law Firm. Le opinioni espresse appartengono all'autrice.