Intelligenza Artificiale, Potere e Responsabilità nell'Era dell'AI Act 2026

Un paper circostanziato tra dati verificabili, rischi sistemici e doveri etici

Salvatore Martino
Riflessione metodologica nell'approccio PENSAI
Agosto 2026

Abstract

Questo paper analizza la traiettoria dell'intelligenza artificiale attraverso dati di mercato verificabili, normativa europea vigente e ricerca accademica di riferimento. L'obiettivo non è né celebrativo né catastrofista: è quello di applicare uno sguardo critico e circostanziato a una tecnologia che, entro il 2026, ha già assunto caratteri di infrastruttura civile globale. Il quadro normativo europeo — l'AI Act (Reg. UE 2024/1689), pienamente operativo dal 2 agosto 2026 nella sua componente di trasparenza — costituisce lo sfondo istituzionale entro cui si colloca ogni riflessione responsabile. Il metodo adottato è coerente con il framework PENSAI (Problem-Evidence-Normalization-Scoring-Action-Intelligence), che privilegia la normalizzazione del problema, la mappatura dell'incertezza e il segnalamento dei red flag prima di qualsiasi raccomandazione operativa.

1. Il Contesto: Dimensioni Reali di un Cambiamento Epocale

1.1 I numeri del mercato

Il mercato globale dell'intelligenza artificiale ha raggiunto circa 248,9 miliardi di dollari nel 2025 [1], con una traiettoria che le principali analisi proiettano verso i 2.673 miliardi entro il 2035, registrando un tasso di crescita annuo composto (CAGR) del 39,73% [2]. Si tratta di cifre che non hanno precedenti nella storia dell'economia digitale, superiori perfino alla curva di crescita di Internet negli anni Novanta.

Nel solo 2024, gli investimenti privati globali in AI hanno toccato 252,3 miliardi di dollari (+26% rispetto al 2023) [3]. Gli Stati Uniti guidano con 109,1 miliardi di dollari, equivalenti a quasi dodici volte gli investimenti cinesi (9,3 miliardi) e ventiquattro volte quelli britannici (4,5 miliardi). Il segmento della AI generativa — modelli in grado di produrre testo, immagini, codice e audio — ha attratto 33,9 miliardi di dollari a livello globale nel 2024, con un incremento del 18,7% rispetto al 2023 e oltre 8,5 volte i livelli del 2022 [3].

L'Europa parte da una posizione di ritardo strutturale: il mercato AI europeo vale oggi meno del 30% di quello statunitense, ma le proiezioni indicano una possibile accelerazione fino a 190 miliardi di euro entro il 2030, anche grazie alle politiche industriali dell'Unione [4].

1.2 Il paradosso dell'adozione reale

A fronte di questi numeri straordinari, l'adozione effettiva nelle imprese rimane sorprendentemente contenuta. I dati BTOS (Business Trends and Outlook Survey) mostrano che, alla metà del 2025, meno del 10% delle aziende nell'economia complessiva utilizza l'AI regolarmente, cifra che sale a poco più del 20% nelle industrie professionali, scientifiche e tecniche [5]. Il 51% delle organizzazioni intervistate da McKinsey nel 2025 ha dichiarato che l'AI generativa stava già riducendo la necessità di posizioni di livello base [6]. Questo dato rivela una tensione strutturale: l'impatto è reale ma diseguale, concentrato in settori specifici e in aziende già attrezzate digitalmente.

2. Il Quadro Normativo: L'AI Act come Spartiacque Storico

2.1 Una timeline che è già storia

Il Regolamento (UE) 2024/1689, noto come AI Act, è entrato in vigore il 1° agosto 2024 e ha avviato una applicazione progressiva che costituisce oggi il principale riferimento globale per la governance dell'intelligenza artificiale [7]. La sua logica è basata sul rischio: quattro livelli (inaccettabile, alto, limitato, minimo) determinano obblighi crescenti.

Le tappe fondamentali sono le seguenti:

• 2 febbraio 2025: entrano in vigore i divieti sulle pratiche AI inaccettabili (manipolazione subliminale, social scoring pubblico, riconoscimento emotivo in contesti sensibili) e l'obbligo di alfabetizzazione AI per le organizzazioni.

• 2 agosto 2025: le norme sui modelli GPAI (General Purpose AI) — GPT-4, Gemini, Claude, DeepSeek — diventano operative per documentazione, diritti d'autore e trasparenza degli dati di addestramento. Soglia tecnica: sistemi addestrati con oltre 10^{23} FLOPS.

• 27 luglio 2026: entra in vigore l'AI Omnibus (Reg. UE 2026/1744), che sposta gli obblighi per i sistemi ad alto rischio dell'Allegato III al 2 dicembre 2027 e quelli dell'Allegato I al 2 agosto 2028, alleggerendo il calendario senza sospendere il Regolamento [8].

• 2 agosto 2026: piena applicazione dell'articolo 50 del Regolamento, che stabilisce gli obblighi di trasparenza verso il pubblico. Da oggi, chi interagisce con un sistema AI deve esserne informato, salvo che ciò sia evidente dal contesto [9]. Le sanzioni per i fornitori di modelli GPAI diventano pienamente operative: fino al 3% del fatturato mondiale o 15 milioni di euro [10].

• 2 dicembre 2026: marcatura leggibile dalle macchine per i contenuti generati da AI; nuovi divieti su deepfake sessuali non consensuali e materiale di abuso su minori.

2.2 Il sistema sanzionatorio

L'AI Act prevede un sistema di sanzioni graduato e severo. Per le violazioni più gravi — pratiche proibite, manipolazione di minori — le ammende arrivano fino a 35 milioni di euro o al 7% del fatturato mondiale annuo. Per i sistemi ad alto rischio, le sanzioni arrivano fino a 15 milioni o al 3% del fatturato. Per informazioni false o fuorvianti fornite alle autorità, il limite scende a 7,5 milioni o all'1% [10]. In Italia, la Legge n. 132/2025 ha affidato la vigilanza all'ACN (Agenzia per la Cybersicurezza Nazionale) e le funzioni di notifica all'AgID [10].

Questo quadro sanzionatorio non è ornamentale. Rappresenta la prima volta nella storia che una grande potenza economica assoggetta i sistemi di intelligenza artificiale a obblighi giuridici cogenti, con un meccanismo di enforcement istituzionalizzato. Il confronto con l'approccio statunitense — ancora fondato prevalentemente su linee guida volontarie — rende l'AI Act un caso unico nel panorama globale della governance tecnologica.

3. I Red Flag: Rischi Sistemici da Non Sottovalutare

3.1 Disinformazione e deepfake: la democrazia sotto pressione

Il World Economic Forum nel Global Risks Report 2024 ha identificato la disinformazione amplificata dall'AI come uno dei rischi sistemici più urgenti a livello globale, potenzialmente capace di compromettere i processi democratici [11]. Il 2024, definito "l'anno più elettorale della storia" con metà della popolazione mondiale chiamata alle urne, ha costituito un banco di prova drammatico.

Le elezioni presidenziali americane del 2024 hanno registrato casi documentati di uso distorto dell'AI: audio falsi attribuiti al Presidente Biden per scoraggiare gli elettori democratici nel New Hampshire; video deepfake condivisi da personalità pubbliche con voce clonata di Kamala Harris [12]. La tecnologia permetteva già nel 2024 deepfake in tempo reale durante videochiamate live, rendendo vulnerabili anche le comunicazioni riservate [12].

> *"L'IA generativa ha creato le condizioni per generare notizie fasulle, immagini finte, audio e video falsi. Non si tratta di un gioco: si altera la correttezza dell'informazione e della comunicazione."*

— Rapporto DIS-AISE sulla sicurezza nazionale e AI, 2025 [13]

Secondo l'Edelman Trust Barometer 2025, la fiducia nelle informazioni online continua a diminuire mentre cresce la percezione di vivere in un ecosistema dominato dalla manipolazione emotiva [14]. Il Reuters Institute Digital News Report conferma che i contenuti ad alto impatto emotivo sono sistematicamente premiati dagli algoritmi delle piattaforme social [14]. Il risultato è una "folla algoritmica" che reagisce prima che i fatti vengano verificati.

3.2 Impatto sul lavoro: tra sostituzione e trasformazione

La ricerca di Goldman Sachs del marzo 2023 — ancora la più citata — stima che l'AI generativa potrebbe automatizzare l'equivalente di 300 milioni di posti di lavoro a tempo pieno a livello globale entro il 2030 [15]. McKinsey stima che l'AI generativa potrebbe automatizzare tra il 60% e il 70% delle attività lavorative in generale entro la stessa data [16]. Il World Economic Forum nel Future of Jobs Report 2025 prevede che 92 milioni di posti di lavoro saranno eliminati entro il 2030 a causa dell'automazione.

Tuttavia, la ricerca economica invita alla prudenza interpretativa. Goldman Sachs ha rivisto le sue proiezioni nel 2025, concludendo che l'impatto sull'occupazione complessiva sarà "lieve e di breve durata" piuttosto che causa di perdite diffuse e permanenti [5]. Questo perché il vero collo di bottiglia non è la capacità tecnologica — McKinsey stima che il 57% delle ore lavorative attuali è già tecnicamente automatizzabile — ma il deployment: la velocità con cui le organizzazioni possono ristrutturare processi, ruoli e competenze [17].

McKinsey proietta che, nel periodo 2016-2030, circa 400 milioni di persone potrebbero essere rimpiazzate dall'AI. Ma nello stesso periodo verrebbero generati tra 555 e 890 milioni di nuovi posti di lavoro [18]. Il saldo atteso è positivo, ma il problema è distributivo: chi perde e chi guadagna non sono necessariamente le stesse persone, né abitano gli stessi territori.

3.3 Bias algoritmici e discriminazione sistemica

Un rischio meno visibile ma strutturalmente grave riguarda i bias insiti nei sistemi AI addestrati su dati storici. Gli algoritmi di AI ereditano — e talvolta amplificano — le discriminazioni presenti nei dati di partenza: razziali, di genere, socioeconomiche. Il 32% dei docenti italiani teme l'amplificazione di bias algoritmici nell'educazione [19]. La Commissione europea ha riconosciuto il problema inserendo nei sistemi ad alto rischio — assunzioni, accesso al credito, valutazioni scolastiche, decisioni giudiziarie — obblighi specifici di sorveglianza umana e gestione degli incidenti.

3.4 La concentrazione del potere

Il mercato AI presenta una concentrazione senza precedenti. Gli investimenti privati statunitensi del 2024 (109,1 miliardi) sono quasi dodici volte quelli cinesi, e l'ecosistema è dominato da un numero ristretto di grandi player: OpenAI, Google DeepMind, Anthropic, Meta AI, Microsoft. Anthropic ha triplicato il fatturato annualizzato passando da 1 a 3 miliardi di dollari nei dodici mesi fino a dicembre 2025 [20]. NVIDIA da sola ha investito 1 miliardo di dollari in 50 partecipazioni in startup AI nel 2024 [20]. Questa concentrazione pone interrogativi seri sulla governance democratica dell'AI: chi decide i valori codificati negli algoritmi? Chi controlla gli errori di sistemi che influenzano milioni di decisioni quotidiane?

4. Opportunità Verificabili: Dove l'AI Crea Valore Reale

Sarebbe disonesto intellettualmente ignorare i benefici documentati. Goldman Sachs stima che l'AI generativa potrebbe incrementare il PIL globale del 7% nei prossimi dieci anni e aumentare la

produttività del lavoro dell'1,5% annuo [15]. Il mercato del PIL è ancora in via di materializzazione, ma in alcuni domini i risultati sono già misurabili:

• Sanità: la FDA statunitense ha autorizzato oltre 110 dispositivi radiologici basati su AI tra gennaio 2024 e maggio 2025, con una riduzione significativa dei tempi diagnostici [20]. Il settore sanitario è proiettato a crescere con il CAGR più elevato tra tutti i settori AI, pari al 38,35% entro il 2031 [20].

• Ricerca scientifica: l'AI ha accelerato la scoperta di nuovi materiali, la progettazione di farmaci e la modellazione climatica in modo documentato e verificabile. AlphaFold di Google DeepMind ha risolto la struttura di oltre 200 milioni di proteine, problema che la biologia molecolare non aveva risolto in cinquant'anni di lavoro.

• Produttività professionale: studi sperimentali (MIT, 2023) mostrano incrementi di produttività del 14-37% per i lavoratori della conoscenza che utilizzano strumenti AI generativa per compiti cognitivi complessi.

• Accessibilità: tecnologie di sintesi vocale, traduzione automatica e supporto cognitivo stanno riducendo barriere per persone con disabilità, con impatti sociali positivi difficilmente quantificabili ma reali.

5. Applicazione del Metodo PENSAI: Analisi della Decisione Collettiva

Il metodo PENSAI — Problem-Evidence-Normalization-Scoring-Action-Intelligence — è un framework decisionale che privilegia la normalizzazione del problema rispetto alla reazione impulsiva, la mappatura dell'incertezza rispetto alla certezza simulata, e il riconoscimento dei red flag prima di qualsiasi raccomandazione operativa.

Applicato al contesto AI 2026, il framework produce il seguente output:

NORMALIZED STATE — Il problema reale

Non siamo di fronte a una scelta binaria tra "adottare" o "rifiutare" l'AI. Siamo di fronte a una transizione tecnologica già in corso, che avviene indipendentemente dalle scelte individuali, e che richiede governance attiva a livello collettivo. Il vero problema non è l'AI in sé, ma la distribuzione asimmetrica dei suoi benefici e dei suoi rischi tra individui, categorie sociali e nazioni.

UNCERTAINTY MAP — Dove non sappiamo

Alta incertezza: impatto occupazionale netto a lungo termine; velocità reale di adozione nelle PMI; efficacia degli strumenti di enforcement dell'AI Act; traiettoria dei modelli beyond-GPT. Media incertezza: concentrazione di mercato e reazione antitrust; capacità delle democrazie di regolare deepfake senza censura. Bassa incertezza: crescita degli investimenti nel breve-medio termine; proliferazione degli strumenti generativi; crescente uso elettorale distorto.

RED FLAGS — Segnali da monitorare

1. La concentrazione del mercato AI in un numero ristretto di attori privati extraeuropei crea una dipendenza strutturale dell'Europa in un settore di sovranità critica.

2. L'alfabetizzazione AI della popolazione e dei decisori pubblici è gravemente insufficiente: solo il 25% dei docenti italiani usa AI nella didattica contro il 37% della media OCSE e il 75% di Singapore [19].

3. I tempi di adeguamento delle competenze lavorative sono significativamente più lenti della velocità di deployment tecnologico, creando un gap potenzialmente destabilizzante.

4. La fiducia istituzionale nell'informazione digitale sta erodendo più rapidamente di quanto i meccanismi di verifica riescano a compensare.

VIA UNICA — La raccomandazione a più alto punteggio

Investire in alfabetizzazione critica, infrastrutture pubbliche di dati e capacità di enforcement regolatorio è la scelta con il rapporto impatto/rischio più elevato nel breve termine. Non è possibile governare ciò che non si comprende. Non è possibile regolare ciò per cui mancano competenze interne alle istituzioni. L'AI Act è una cornice necessaria ma non sufficiente: il valore reale della norma dipende dalla capacità delle autorità nazionali di applicarla con cognizione tecnica.

NON FARE — Le trappole da evitare

• Non trattare l'AI Act come un adempimento burocratico: è un'occasione storica di ridefinire il patto tra tecnologia e democrazia.

• Non cedere al determinismo tecnologico: la tecnologia non ha un destino, ha traiettorie che dipendono da scelte politiche ed economiche.

• Non confondere la velocità con l'urgenza: decisioni rapide in contesti di alta incertezza producono spesso danni superiori a quelli che intendono prevenire.

6. Conclusioni

L'intelligenza artificiale è già infrastruttura: permea sistemi sanitari, giudiziari, finanziari, informativi e produttivi in modo che non ammette più reversibilità semplice. Il mercato da 248 miliardi di dollari del 2025, destinato a decuplicare entro il 2035, non è una proiezione speculativa: è la somma di decisioni di investimento già prese e contratti già firmati.

L'AI Act europeo, con la sua piena operatività degli obblighi di trasparenza a partire dal 2 agosto 2026, è il contributo più significativo che le democrazie liberali abbiano prodotto finora alla governance globale dell'AI. Non è perfetto — i rinvii introdotti dall'AI Omnibus riflettono la difficoltà politica di applicare standard elevati in un contesto di competizione globale — ma è il primo tentativo serio di assoggettare il potere algoritmico al diritto democratico.

Ciò che questa riflessione vorrebbe suggerire, con il metodo PENSAI e nella consapevolezza dei limiti propri di ogni analisi, è che la domanda cruciale non è "l'AI è buona o cattiva?" ma "chi decide cosa fa, per chi e con quali conseguenze?". Questa è una domanda politica prima che tecnologica. E le risposte, per essere legittime, devono essere trasparenti, contestabili e democraticamente responsabili — esattamente ciò che l'articolo 50 dell'AI Act, operativo da oggi, esige per ogni sistema che interagisce con un essere umano.

Bibliografia e Fonti di Riferimento

Le seguenti fonti sono tutte verificabili e di pubblico dominio al 2 agosto 2026.

[1] The Insight Partners (2025). "AI Market Size and Forecast 2021-2034". Disponibile su: theinsightpartners.com/reports/artificial-intelligence-market

[2] Business Research Insights (2026). "Artificial Intelligence (AI) Market — Global Forecast to 2035". Aggiornato: 29 giugno 2026.

[3] Stanford University Human-Centered AI (HAI). "Artificial Intelligence Index Report 2025". AI4Business, 28 aprile 2025. Disponibile su: ai4business.it

[4] Thunderbit (2026). "Le 150 Statistiche Più Importanti sull'Intelligenza Artificiale per il 2026". Aggiornato: febbraio 2026.

[5] Goldman Sachs Research (2025). "AI and the Labor Market". Citato in: PB Consulting (settembre 2025). "L'intelligenza artificiale e il futuro del lavoro". consultingpb.com

[6] McKinsey & Company (2025). "The State of AI 2025". Citato in: Forbes Italia, 26 maggio 2026.

[7] Commissione Europea (2024). Regolamento (UE) 2024/1689 — AI Act. Gazzetta Ufficiale dell'Unione Europea, 12 luglio 2024. digital-strategy.ec.europa.eu

[8] Federprivacy (luglio 2026). "AI Omnibus in vigore dal 27 luglio 2026". federprivacy.org. Reg. UE 2026/1744, GUUE 24 luglio 2026.

[9] Il Sole 24 Ore (31 luglio 2026). "AI Act europeo, ecco gli obblighi già in vigore per imprese e professionisti". ilsole24ore.com

[10] AgendaDigitale.eu (agosto 2026). "AI Act, cosa entra in vigore dal 2 agosto 2026 e cosa slitta". agendadigitale.eu; PMI.it (luglio 2026). "Intelligenza artificiale: dal 2 agosto nuove regole UE". pmi.it

[11] World Economic Forum (2024). "Global Risks Report 2024". Weforum.org. Citato in: Filodiritto, 2024.

[12] Ecostampa (novembre 2025). "Disinformazione e AI: il fact checking contro fake news e manipolazioni". ecostampa.it

[13] ICT Security Magazine (marzo 2025). "Intelligenza artificiale e sicurezza nazionale 2025". Rapporto DIS-AISE. ictsecuritymagazine.com

[14] AgendaDigitale.eu (giugno 2026). "Difendere la reputazione quando deepfake e algoritmi accelerano le crisi". Citazione: Edelman Trust Barometer 2025; Reuters Institute Digital News Report. agendadigitale.eu

[15] Goldman Sachs Economics Research (marzo 2023). "The Potentially Large Effects of Artificial Intelligence on Economic Growth". Citato in: myp.srl, ottobre 2025.

[16] McKinsey Global Institute (2023). "The Economic Potential of Generative AI". Citato in: Prof. Hung-Yi Chen (febbraio 2026). "AI Job Displacement 2026". hungyichen.com

[17] AImultiple (luglio 2026). "Oltre 20 previsioni degli esperti sulla perdita di posti di lavoro dovuta all'IA". aimultiple.com/it/ai-job-loss

[18] Digital4PRO (febbraio 2025). "AI: L'impatto sul lavoro e sull'occupazione". McKinsey Global Institute citato. digital4pro.com

[19] OCSE e Fondazione Agnelli (dicembre 2025). "AI Adoption in the Education System". Indagine TALIS OCSE 2024. Citato in: Orizzonti Politici, maggio 2026.

[20] Mordor Intelligence (2026). "Dimensioni del mercato dell'intelligenza artificiale". mordorintelligence.it. Dati FDA: gennaio 2024 – maggio 2025.

⚖️ Nota metodologica e dichiarazione d'autore

Questo paper è una riflessione elaborata da Salvatore Martino, che ha scelto di applicare il proprio metodo PENSAI, all'analisi del cambiamento tecnologico in atto. Il metodo PENSAI privilegia la normalizzazione del problema rispetto alla reazione emotiva, la mappatura esplicita dell'incertezza, il riconoscimento dei segnali deboli e l'identificazione dei red flag prima di qualsiasi raccomandazione operativa. Non è un metodo che garantisce risposte certe: è un metodo che garantisce domande oneste.

La riflessione è coerente — nella misura in cui è umanamente possibile — con i principi dell'AI Act 2026, in particolare con gli obblighi di trasparenza (art. 50 Reg. UE 2024/1689), pienamente vigenti dal 2 agosto 2026. Ciò significa: dichiarare che questo testo è stato elaborato con il supporto di strumenti di intelligenza artificiale (Claude, Anthropic); che le fonti citate sono verificabili e di dominio pubblico; che nessuna affermazione è presentata con certezza superiore a quella che i dati disponibili consentono; che l'autore umano mantiene la responsabilità editoriale e intellettuale piena del contenuto.

Salvatore Martino — Agosto 2026