



UNIVERSITÀ DEGLI STUDI DI MILANO  
DIPARTIMENTO DI DIRITTO PUBBLICO  
ITALIANO E SOVRANAZIONALE



# PoliS AI NEWS

Newsletter sull'Intelligenza Artificiale  
a cura di PoliS-Lombardia

Anno III – n. 3/2026

In questo numero

In evidenza

Focus

Normativa

Applicazioni alla Pubblica Amministrazione

AI in pillole

Notizie

Commenti

Corsi, convegni e pubblicazioni

## In questo numero

Che rapporto c'è tra la diffusione dell'Intelligenza artificiale e la salute delle istituzioni? Ne parla un articolo di due giuristi americani che sta facendo molto discutere, perché trae conclusioni nette e decisamente a tinte

fosche... Lo trovate tra le notizie “In evidenza”, assieme alla nuova “Costituzione” di Anthropic, che punta a fissare delle (auto)regole per l’universo AI, ad una panoramica su AI e mondo del lavoro realizzata dal World Economic Forum e alla segnalazione di un nuovo soggetto che emerge nella geopolitica dell’Intelligenza artificiale: gli Emirati Arabi. Il “Focus” è sul conflitto aperto tra Commissione europea e XCorp., la galassia di Elon Musk. In “Normativa”, un’analisi sull’uso militare dell’AI (molto differente tra Usa e Europa), mentre la “pillola didattica” è dedicata alla valutazione d’impatto nell’uso dell’AI. Poi, come sempre, notizie, commenti, applicazioni alla Pubblica amministrazione, strumenti di lavoro. Buona lettura!

## In evidenza

### Ma l’AI fa bene o no alle istituzioni?

L’Intelligenza artificiale è davvero solo un motore che ci rende più veloci e performanti? O il tecno-ottimismo dominante sta oscurando rischi più profondi, che non riguardano solo gli abusi ma la natura stessa di queste tecnologie? Su questo punto, il confronto tra studiosi, policy-maker e professionisti è più acceso che mai. Un contributo recente al dibattito arriva da due professori di Boston, Woodrow Hartzog e Jessica Silbey, autori del paper [How AI Destroys Institutions](#). La loro tesi è radicale: l’AI non è affatto uno strumento neutrale da aggiungere alla cassetta degli attrezzi sociali. Anzi, è strutturalmente incompatibile con la tenuta delle istituzioni civiche democratiche.

Quando parliamo di istituzioni non ci riferiamo a edifici, strutture burocratiche o organigrammi, ma alla “spina dorsale invisibile della vita sociale”: quell’intreccio di norme, pratiche e responsabilità che converte il caos in ordine condiviso. Ne sono un esempio l’istruzione superiore, la medicina, il diritto o la libertà di espressione. Università, ospedali, tribunali e redazioni (le cosiddette “organizzazioni”) ne sono il corpo fisico. Molte di queste istituzioni sono “pronte per essere ottimizzate dall’AI” ed è qui che inizia la frizione.

Il problema, spiegano gli studiosi, risiede nelle *affordances*: le caratteristiche intrinseche di una tecnologia che orientano il modo in cui può essere usata. Il punto è che le *affordances* dell’AI non sono neutre. Quando entrano in contatto con il tessuto istituzionale, producono tre effetti destabilizzanti.

Il primo riguarda l’expertise. L’AI induce a delegare compiti cognitivi a sistemi che simulano competenza senza possederla, creando una dipendenza che indebolisce le capacità umane. Quando il sistema funziona, le persone si “atrofizzano”; quando sbaglia, occorre investire tempo ed energie per correggere gli errori. In entrambi i casi, la struttura istituzionale viene compromessa.

Il secondo effetto è il cortocircuito del processo decisionale, che si verifica quando scelte morali e giudizi complessi vengono di fatto delegati agli sviluppatori di AI, tagliando fuori quei momenti di riflessione e confronto che, insieme all’assunzione di responsabilità, permettono alle istituzioni di funzionare. Un’altra criticità è che l’AI “decide” seguendo logiche di ottimizzazione: rifiuta rimborsi sanitari, assegna priorità, filtra candidature, senza offrire ragioni pubbliche. Ma più queste logiche diventano opache, più si assottiglia la legittimità che le istituzioni hanno costruito nel tempo.

Il terzo meccanismo è l’isolamento sociale. Riducendo interazioni e filtrando il confronto, l’AI erode la capacità di costruire obiettivi condivisi, gestire conflitti e sviluppare empatia. Meno relazioni significa meno coesione, e una società senza coesione, osservano gli autori, non reggerà l’urto della trasformazione tecnologica.

Le vittime di questi processi? Lo stato di diritto, dove prevedibilità e trasparenza delle procedure cedono il passo ad automatismi indecifrabili. L’istruzione superiore, dove l’AI appiattisce l’apprendimento, genera

contenuti mediocri e indirizza la ricerca solo verso ciò che è misurabile. Il **giornalismo**, soffocato da contenuti a basso costo che indeboliscono l'autorevolezza delle fonti e producono notizie sempre più uniformi e disconnesse dai lettori. **La vita civica**, dove erosione della fiducia e isolamento prosciugano quel capitale relazionale che permette alle persone di unirsi per affrontare problemi comuni.

Nelle loro conclusioni, Hartzog e Silbey **evitano soluzioni semplicistiche**. Avvertono che serviranno **interventi sulle "vulnerabilità strutturali"** – dalle disuguaglianze ai processi democratici indeboliti, come il voto – e **azioni radicate nel locale** perché "accadono cose positive quando le persone pensano e agiscono a quel livello". Su un punto non esitano: **non saranno i codici etici autoregolatori a fare da argine**. Incentivi economici e pressioni operative sono spesso troppo forti per resistere alle tentazioni dell'AI.

C'è però un **filo che attraversa i tre meccanismi distruttivi dell'AI** descritti dagli autori: **la scomparsa della mediazione**, quella capacità di trasmettere saperi, discutere, negoziare, entrare in conflitto costruttivo. Il risultato è uno slittamento quasi impercettibile: da una società che vive di interpretazioni e negoziazioni a una che si affida a sistemi standardizzati, più "lisci" ma ciechi. Troppo pessimismo? Il dibattito è aperto...

\*Il [testo del paper](#)

### ***Anthropic adotta una Costituzione per Claude (ma che cosa rimane dello human in the loop?)***

È di alcuni giorni fa la notizia della **pubblicazione, da parte di Anthropic, della [Claude's new Constitution](#)**, documento che aggiorna le precedenti versioni (già nel 2023 Anthropic aveva abbozzato un testo) e che definisce i principi etici e operativi alla base del suo modello AI.

L'obiettivo è molto ambizioso: costruire un'AI utile, sicura, che sia coerente con una serie di principi ispirati ai diritti umani, alla sicurezza e alla responsabilità tecnologica, **senza dipendere esclusivamente dal giudizio esterno di valutatori umani**.

A dispetto del nome, non è una costituzione in senso giuridico, ma piuttosto un **modello di addestramento alternativo** rispetto ai tradizionali sistemi basati sul giudizio umano, che prende il nome di **"Constitutional AI"**: invece di affidarsi solo a valutatori esterni, **Claude viene guidato da un insieme di regole scritte che permettono al modello di auto-correggersi**.

Il documento **stabilisce le regole che devono orientare il modello a generare i testi e a gestire le richieste degli utenti**. In particolare, deve evitare contenuti violenti, discriminatori o potenzialmente dannosi, proteggendo la privacy degli utenti. Vengono definiti anche **limiti precisi: nessuna presa di posizione politica, nessuna manipolazione dell'utente, nessun consiglio medico o legale** personalizzato.

Allo stesso tempo, il modello è incoraggiato a mantenere nei confronti degli utenti un tono collaborativo, chiaro e rispettoso, privilegiando risposte utili e comprensibili e, qualora una richiesta superi i suoi limiti, deve dichiararlo apertamente e suggerire di rivolgersi a professionisti qualificati. Uno dei punti chiave della Claude's Constitution è proprio lo **scopo benevolo**, lecito ed etico nell'uso del modello: **l'AI deve agire nell'interesse dell'utente**, mostrando empatia e prudenza, soprattutto quando le domande riguardano temi sensibili o potenzialmente rischiosi.

Con questo documento, Anthropic punta a distinguersi nel panorama dell'AI generativa, proponendo un modello di AI che opera all'interno di una struttura etica dichiarata e verificabile. Ma resta **un dubbio: se il modello è in grado di autocorreggersi, quale spazio rimane per il controllo umano?**

\*Il [testo del documento](#)

Per approfondire:

[M. Carmignani e A. Longo, La Costituzione di Claude: alla ricerca di un'AI etica | Agenda Digitale, 28 gennaio 2026](#)

## L'AI e il lavoro che ci aspetta nel 2030

Mentre l'Intelligenza artificiale si afferma come strumento sempre più usato dalle aziende (sono l'88% quelle che ormai la utilizzano almeno in una funzione), e mentre l'*Harvard Business Review* [mette in dubbio che l'AI alleggerisca il lavoro](#) (in realtà lo rende molto più complesso e pervasivo, almeno da una ricerca sul campo dei ricercatori di Yale e Berkley), gli analisti del **World Economic Forum (Wef)** si chiedono **quale futuro si stia preparando per il mercato del lavoro**. Interrogativo a cui risulta difficile rispondere, a causa dei rapidi cambiamenti in atto, ma anche perché il futuro si prepara con le scelte strategiche di oggi.

Un'indagine svolta fra oltre diecimila dirigenti in tutto il mondo attesta che per **il 54% degli intervistati l'AI toglierà posti di lavoro esistenti, mentre per il 24% ne creerà di nuovi** e molti. Per il 40% i profitti delle aziende cresceranno e migliorerà l'accessibilità a beni e servizi (37%), anche dal punto di vista economico (37%). C'è anche **chi teme l'aggravarsi della discriminazione nei confronti di specifici gruppi** demografici.

Il libro bianco [Four Futures for Jobs in the New Economy: AI and Talent in 2030](#) prende in esame **quattro possibili scenari futuri**, in cui il progresso delle competenze digitali a disposizione risulterebbe determinante rispetto al successo delle aziende, ma anche agli equilibri sociali.

**Il primo scenario, definito accelerato**, vedrebbe una rapida affermazione dell'AI, che moltiplicherebbe la capacità produttiva, cancellando posti di lavoro e quindi aumentando la disoccupazione rispetto al 5% del 2025, pur creandone di nuovi. Welfare, etica e governance faticerebbero a tenere il passo delle trasformazioni.

**Il secondo scenario, o "era della sostituzione"**, sarebbe contraddistinto dall'incapacità dei lavoratori di formarsi e riqualificarsi al ritmo del progresso dell'AI. I maggiori profitti delle aziende avrebbero un prezzo: picchi di disoccupazione e disgregazione sociale, con una perdita di fiducia da parte dei consumatori (rispetto al 48,8%).

**L'economia del copilota, il terzo scenario**, sarebbe caratterizzato da un progresso più graduale sia dell'AI sia delle competenze adatte a gestirla, grazie all'investimento di governi e imprese in formazione, mobilità, infrastrutture digitali e governance dell'Intelligenza artificiale.

Infine, si potrebbe verificare anche lo **scenario "progressi in stallo"**, in cui al progresso dell'AI non corrisponderebbe quello dei lavoratori. Si creerebbero così divari fra le aziende "competenti", in grado di arricchirsi, e quelle che non riescono a stare al passo, impoverendosi.

Quale scenario si realizzerà **dipenderà dalle strategie e dagli investimenti che si decidono oggi**. Il documento sottolinea nel dettaglio potenziali benefici e rischi, con le strategie che sarebbe opportuno assumere.

Sulla base di una serie di workshop e consultazioni con responsabili della strategia ed esperti, il Wef riporta anche alcune azioni ritenute da tutti positive: **iniziare in piccolo, costruire rapidamente, scalare ciò che funziona**. In particolare viene attribuita importanza a cinque fattori: l'apprendimento dell'AI nel flusso di lavoro, così che le persone possano acquisire competenze personalizzate e specifiche; gli investimenti nella governance dei dati e nell'infrastruttura; il sostegno alla cultura organizzativa e alla fiducia nelle tecnologie emergenti, senza trascurare la creazione di team intergenerazionali, per evitare divari culturali e, infine, la cura di partnership strategiche.

\*[Il libro bianco del World Economic Forum](#)

## E se il Terzo Polo dell'AI fosse in Medio Oriente?

Gli Emirati Arabi Uniti hanno lanciato [K2 Think](#), un **modello di AI completamente opensource** sviluppato dalla **Mohamed bin Zayed University of Artificial Intelligence** di Abu Dhabi.

Il sistema si distingue per un livello di trasparenza abbastanza raro nel settore: sono **pubblici la documentazione completa, i dati di addestramento e il codice sorgente**. Una scelta significativa, in un momento storico in cui molte Big Tech occidentali tendono a non dare accesso ai dettagli dei propri sistemi. E anche i modelli openweight cinesi – come quelli di Alibaba o DeepSeek –, pur essendo parzialmente aperti, non raggiungono il livello di trasparenza del competitor mediorientale. Secondo le prime valutazioni degli utenti, K2 Think compete con i modelli aperti di Nvidia, OpenAI e Alibaba, mostrando prestazioni solide soprattutto nella **riduzione degli errori e delle cosiddette allucinazioni**, un aspetto cruciale per applicazioni in ambito pubblico e industriale.

Ma la mossa è solo una parte della strategia digitale degli Emirati. [Abu Dhabi sta investendo anche in progetti occidentali](#), come [Stargate](#) (coguidato da OpenAI), e **supporta partnership infrastrutturali** con aziende come Microsoft e BlackRock.

Questa strategia duplice – costruire capacità nazionali e allo stesso tempo partecipare ai grandi ecosistemi internazionali dell'AI – riflette **l'ambizione degli Emirati di diventare un attore globale nel settore**, non solo come consumatore ma anche produttore e partner tecnologico. E se fosse la via per il Terzo Polo dell'AI, dopo Usa e Cina?

## Focus



## La Commissione europea apre un'istruttoria sui sistemi di raccomandazione di Grok e X a cura di Marco Bassini – Tilburg University

Lo scorso 25 gennaio, la Commissione europea ha annunciato una nuova indagine nei confronti di X Corp. (la società di Elon Musk che gestisce la piattaforma X e il modello di AI Grok) con l'obiettivo di **verificare la conformità dei loro sistemi di raccomandazione automatica alle norme del Digital Services Act (DSA)**, il quadro giuridico adottato nel 2022 dall'Ue per disciplinare i servizi digitali e governare i rischi sistemici che possono generare.

L'iniziativa ha fatto seguito alle preoccupazioni sull'operato degli algoritmi che suggeriscono contenuti agli utenti sulle principali piattaforme online. Nello specifico, **la Commissione intende accertare se i meccanismi di raccomandazione di X e di Grok**, che possono influenzare fortemente il consumo di informazione e l'esperienza degli utenti, **rispettino i requisiti di gestione dei rischi sistemici e di tutela dei diritti fondamentali imposti dal DSA**.

Tra le **aree di rischio** esaminate figurano la **diffusione di contenuti illegali o dannosi**, laddove i sistemi di raccomandazione possono amplificare la visibilità di contenuti problematici. Ai sensi del DSA, **i sistemi di AI sono soggetti al quadro di gestione dei rischi se sono integrati in piattaforme online, o motori di ricerca, di dimensioni molto grandi**.

Nel caso di **Grok, l'integrazione nella piattaforma X** (che è stata designata come piattaforma online di dimensioni molto grandi) **determina l'obbligo** per il relativo fornitore del servizio di **effettuare una verifica sull'esistenza di possibili rischi sistemici** (come quelli derivanti dalla diffusione di contenuti illeciti, fra cui rientrano senz'altro i *deepfakes*) e di **adottare le necessarie misure di mitigazione**. Tale apprezzamento deve **tenere in conto anche il ruolo e il contributo di eventuali sistemi di raccomandazione adottati dalle piattaforme**.

La Commissione, così, intende **verificare se il gestore della piattaforma X abbia correttamente valutato i rischi sistemici di diffusione di contenuti espliciti e adottato i necessari rimedi**, alla luce delle caratteristiche tecniche dell'AI integrata.

L'indagine della Commissione si inserisce in un **contesto di crescente attenzione da parte dell'Ue nei confronti delle grandi piattaforme digitali**, molto spesso di origine statunitense. Negli ultimi anni l'Europa ha intensificato gli sforzi per garantire l'applicazione del DSA e di altri strumenti legislativi (come il Digital Markets Act) per **incentivare la responsabilizzazione degli operatori di servizi su scala globale**, soprattutto alla luce della diffusione dell'AI che è sempre più integrata nei servizi digitali ponendo seri problemi in termini di trasparenza.

In particolare, i procedimenti della Commissione prendono di mira quelle tecnologie che pongono **rischi di particolare entità rispetto alla tutela di alcuni diritti fondamentali**, tra cui la privacy e la libertà di espressione degli utenti. In base al DSA, **le autorità possono avviare istruttorie formali quando emergono elementi significativi di violazioni delle norme** che impongono obblighi stringenti di mitigazione dei rischi e di trasparenza. L'esito di questa indagine potrebbe avere **effetti rilevanti sulla regolazione dell'AI e dei servizi digitali**, rafforzando il potere di supervisione dell'Ue sul modo in cui le piattaforme modellano il consumo di contenuti e affrontano rischi sistemici.

[\\*Per approfondire](#)

# Normativa

L'uso dell'AI nelle strategie militari, dal memorandum del Department of War USA alle regole Ue

a cura di Annalisa Negrelli

**Il Department of War USA (l'ex Ministero della Difesa) accelera sulla trasformazione digitale del settore militare e mira a diventare una forza "AI-first".** È quanto emerge nel **memorandum del 9 gennaio 2026**: la nuova strategia sull'AI firmata dal Segretario alla Guerra Pete Hegseth, in linea con l'executive order del presidente Trump per consolidare la supremazia americana nelle tecnologie emergenti.

L'AI, si legge nel documento, «ridisegnerà la natura dei conflitti nel prossimo decennio» in quella che viene chiamata la corsa globale agli armamenti e alla «dominanza militare AI».

La strategia introduce **sette Pace-Setting Projects**, suddivisi in tre aree: **operazioni militari, intelligence e gestione interna**. Tra questi, figurano programmi per sperimentare tattiche basate su sciami autonomi, sviluppare agenti intelligenti per la gestione del campo di battaglia e il supporto decisionale, potenziare le capacità di simulazione avanzata.

Altri progetti puntano a **trasformare l'intelligence in «capacità operative concrete» (armi)** e a sviluppare strumenti di genAI fruibili direttamente dai tre milioni di dipendenti civili e militari (GenAI.mil).

Il memorandum **segna un cambio di passo culturale: la velocità diventa un criterio operativo («speed wins»)**. In tal senso, il documento auspica la rapida rimozione di ostacoli burocratici e promuove un approccio "AI-native" nella revisione di tattiche, procedure e concetti operativi. Infine, la strategia prende le distanze da modelli con filtri ideologici, chiedendo sistemi «oggettivi» e utilizzabili per «qualsiasi scopo lecito».

Il **2026** viene indicato dal documento come **l'anno in cui gli USA intendono accrescere in modo decisivo gli standard della dominanza militare AI** rispetto al resto del mondo.

Di questo approccio, tuttavia, si intravedono già **le criticità nei conflitti attuali**, come l'uso di [robot](#) e [droni killer nella guerra Russia-Ucraina](#), e di *lethal autonomous weapon systems*, sui quali le [Nazioni Unite si sono più volte espresse](#), lanciando un monito sui potenziali rischi.

L'utilizzo dell'AI in guerra pone non poche **questioni etiche**, tra cui la responsabilità, il rispetto del diritto umanitario internazionale e il rischio di errori o, peggio, di una pericolosa escalation dei conflitti dovuta a una diminuzione della supervisione umana.

**In assenza di un quadro internazionale unificato**, il dibattito globale sulla regolamentazione dell'AI militare si è intensificato, con **approcci contrastanti tra gli USA** – che, come si legge nel memorandum, promuovono standard flessibili e favorevoli all'innovazione – e **l'Ue** che adotta un modello centrato sull'uomo e basato sul rischio, secondo quanto previsto dall'**AI Act** che, di fatto, **ne esclude l'uso militare**.

Il **Parlamento europeo riconosce l'importanza strategica dell'AI nella difesa, ma chiede regolamentazione e il divieto delle armi autonome letali**. La Commissione Speciale del Parlamento sull'Intelligenza artificiale nell'era digitale sottolinea la necessità di linee guida etiche sull'AI nella difesa, chiedendo norme internazionali sui *lethal autonomous weapon systems*, misure di sicurezza informatica robuste e cooperazione globale nella regolamentazione dell'AI militare.

Per approfondire:

[Artificial Intelligence Strategy for the Department of War | Secretary of War, 9 gennaio 2026](#)

[C. Alberico, Dalla difesa militare al "Democratic use": la nuova sicurezza europea | Agenda Digitale, 25 novembre 2025](#)

[Istituto di Ricerche Internazionali Archivio Disarmo, Unione degli scienziati per il disarmo, Lo stato dell'Intelligenza Artificiale in ambito militare e le prospettive di regolazione a livello nazionale, europeo e internazionale | MAECI, settembre 2025](#)

# Applicazioni alla Pubblica Amministrazione

## ITALIA

### Olympics Gpt – Milano Cortina 2026

Per Milano Cortina 2026, è stata creata un'AI ad hoc: **Olympic Gpt**, specializzata nella produzione di contenuti per il sito [Olympics.com](https://olympics.com), in cui reperire varie informazioni, dai risultati in tempo reale ai regolamenti sportivi.

[M. Fiordalisi, È il giorno delle Olimpiadi di Milano Cortina 2026, ecco tutte le tecnologie che vedremo per la prima volta | Wired, 6 febbraio 2026](#)

### AI PACT

Finanziato dal PNRR, promosso dal MIMIT e coordinato dalla SDA Bocconi School of Management, con il coinvolgimento di Fondazione IFEL ANCI, il progetto mette a disposizione delle PA servizi specialistici per favorire un'adozione consapevole dell'AI nei processi amministrativi, migliorando l'organizzazione interna e la qualità dei servizi.

[AI-PACT: quando l'innovazione diventa uno strumento per migliorare i servizi pubblici | Città Metropolitana di Genova](#)

### ZENiA – Comune di Genova

Assistente virtuale multilingue, attivo h24/7, per semplificare il rapporto tra Amministrazione e cittadini, consente a residenti, utenti e visitatori di orientarsi con facilità tra servizi, uffici, informazioni, e con funzioni inclusive e accessibili per tutti.

[ZENiA | Comune di Genova](#)

## UNIONE EUROPEA

### ALIA – Spagna

Iniziativa pionieristica nell'Ue che mira a fornire un'infrastruttura pubblica di risorse AI, tra cui modelli linguistici aperti e trasparenti, per promuovere lo spagnolo e le lingue co-ufficiali (catalano e valenciano, basco e galiziano) nello sviluppo e nell'implementazione dell'AI nel mondo.

[ALIA | English](#)

## MONDO

### GenAI.mil

Piattaforma sviluppata dal Dipartimento della Guerra USA per permettere a migliaia di militari di usare Gemini, Claude, Grok e ChatGPT sulle reti governative, a livello di classificazione IL5.

[The War Department Unleashes AI on New GenAI.mil Platform | U.S. Department of War, 9 dicembre 2025](#)

## AI in pillole

La Valutazione d'impatto del sistema di AI: obiettivi e obblighi normativi.  
a cura di Annalisa Negrelli

Per **AI System Impact Assessment (AIIA)** si intende quel processo formale di natura iterativa che accompagna l'intero ciclo di vita dei sistemi AI, dalla fase di progettazione fino all'implementazione operativa, inteso a **identificare, valutare e affrontare l'impatto su individui, gruppi e società derivante da prodotti o servizi che impiegano AI**. Esso si configura come uno strumento metodologico essenziale per promuovere una governance responsabile dei sistemi AI, integrando aspetti etico-sociali (equità, diritti fondamentali), tecnici (accuratezza, robustezza), organizzativi (competenze, sorveglianza umana) e giuridici (basi legali, proporzionalità). L'AI Impact Assessment utilizza infatti quali parametri di misurazione alcuni principi condivisi a livello internazionale in tema di *"safe and trustworthy AI"*, tra cui *accountability and transparency; fairness, safety, security and resilience; explainability and interpretability; validity and reliability; privacy*.

In un contesto come quello attuale, in cui l'attenzione verso i potenziali rischi dell'AI è aumentata di pari passo con l'espansione delle sue applicazioni, l'AIIA **consente di perseguire obiettivi, eseguire controlli, adottare decisioni di rilascio, monitoraggio e miglioramento continuo**, garantendo la conformità ai principi etici e normativi riconosciuti a livello internazionale. Al contempo essa concorre a **individuare, analizzare e mitigare i rischi associati all'adozione dell'AI e i potenziali impatti negativi sulla società, l'economia e i diritti individuali**, ottemperando ai doveri informativi verso autorità e stakeholder.

#### *La diffusione dell'AIIA in Italia e nel mondo*

A questo proposito, sotto il profilo normativo, l'EU AI Act (Regolamento Ue 2024/1689) ha introdotto l'obbligo di valutazioni d'impatto per i sistemi classificati come "ad alto rischio", sul presupposto che *bias*, errori o usi impropri dell'AI possano compromettere diritti fondamentali e processi decisionali democratici.

L'obbligo di effettuare una **Valutazione di Impatto sui Diritti Fondamentali, la cosiddetta FRIA (Fundamental Rights Impact Assessment)**, introdotto dall'art. 29, **entra in vigore per i sistemi di AI ad alto rischio a partire dal 2 agosto 2026**. Questo strumento – di controllo preventivo – ha lo specifico obiettivo di assicurare tre beni pubblici fondamentali: a) **l'omogeneità della regolazione e della prevenzione dei rischi** a livello europeo per tutti gli usi di sistemi ad alto rischio; b) **la fiducia dei cittadini** nelle azioni di supporto, diffusione e utilizzo di questi ultimi; c) **la credibilità delle istituzioni regolative**.

Analogo presupposto di tutela dei diritti fondamentali emerge nella Convenzione Quadro del Consiglio d'Europa relativa all'impatto dell'AI su diritti umani, democrazia e stato di diritto, con l'adozione della **HUDERIA (Human Rights Democracy and Rule of Law Impact Assessment)**, strumento che può essere utilizzato da attori pubblici e privati per contribuire a identificare e affrontare i rischi e gli impatti sui diritti umani, sulla democrazia e sullo Stato di diritto, in una fase successiva, **durante l'intero ciclo di vita dei sistemi di AI**.

In Europa, diversi Paesi hanno già adottato framework specifici per l'AIIA, come rileva il **Joint Research Centre (JRC) della Commissione Europea** – servizio scientifico che produce analisi, studi e linee guida, **tra cui proprio quelle sull'AI Impact Assessment**. Il rapporto *"AI WATCH: Artificial Intelligence for the public sector"* richiama, quali esempi "virtuosi", le strategie di *AI Governance* messe in atto da [Spagna](#) e [Norvegia](#). Sono del 2025 le linee guida AGID sull'AI Impact Assessment in Italia.

Analogamente, a livello internazionale, nel documento *"State of implementation of the OECD AI Principles"*, l'OCSE evidenzia come diversi Paesi nel mondo stiano già adottando framework specifici per l'AIIA. Esempi? Il Canada, che ha sviluppato uno [strumento di valutazione dell'impatto algoritmico, sotto forma di questionario per il settore pubblico e le agenzie governative](#). O il Regno Unito, che ha introdotto delle [linee guida per l'uso etico dell'AI nel settore pubblico](#).

L'AIIA è un elemento cardine per conciliare innovazione tecnologica e tutela dei diritti fondamentali. La sfida futura consiste nella **standardizzazione delle metodologie** utilizzate per la misurazione dell'impatto (mantenendo al contempo la flessibilità necessaria per adattarsi all'evoluzione tecnologica), oltre che nella **promozione di una cultura della valutazione d'impatto** che non si riduca a mero adempimento burocratico, ma favorisca piuttosto la fiducia degli utenti.

Per approfondire:

[O. Pollicino, DPIA e FRIA: guida completa alla valutazione del rischio nell'AI | Agenda Digitale, 10 novembre 2025](#)  
[Valutazione d'impatto dell'Intelligenza artificiale | AGID, 14 febbraio 2025](#)

## Notizie

[K. Conger e R. Mac, Elon Musk Wants to Build an A.I. Satellite Factory on the Moon | The New York Times, 10 febbraio 2026](#)

[T. Brambilla, Scontro Ue-Meta, Bruxelles pronta a misure se l'azienda non cambierà le politiche di WhatsApp sull'AI | AI News, 9 febbraio 2026](#)

[Safer Internet Day, il monito di Telefono Azzurro: l'AI usata dal 75% dei ragazzi italiani, servono regole contro i rischi | Corriere della Sera, 9 febbraio 2026](#)

[G. Di Feo, Robot, licenza di uccidere illimitata: nessuna moratoria sulle macchine assassine | La Repubblica, 8 febbraio 2026](#)

[Adolescente maltrattato per i voti usa l'intelligenza artificiale per chiedere aiuto a Novara | Il Sole 24 Ore, 6 febbraio 2026](#)

[P. L. Pisa, Scontri a Torino, la foto degli agenti aggrediti è IA. Polizia: "Preso dal web, non è opera nostra" | La Repubblica, 6 febbraio 2026](#)

[U. Torelli, Generazione Ai, giovani più connessi ma soli, nove su 10 usano gli algoritmi | Corriere della Sera, 5 febbraio 2025](#)

[P. Arosio, Marte, ora è l'intelligenza artificiale a guidare l'esplorazione | La Repubblica, 4 febbraio 2026](#)

[V. Leone, L'intelligenza artificiale che vede la guerra, come un software ucraino individua armi e soldati in pochi secondi | Wired, 4 febbraio 2026](#)

[A. Magnani, Dalla prevenzione ai vaccini, come l'Intelligenza artificiale può aiutare la sanità africana | Il Sole 24 Ore, 3 febbraio 2026](#)

[P. L. Pisa, Moltbook, il social per le IA: cos'è e come funziona per noi umani | La Repubblica, 2 febbraio 2026](#)

[A. Greenberg, Giocattoli AI, un'azienda ha lasciato migliaia di chat di bambini alla portata di chiunque | Wired, 2 febbraio 2026](#)

## Commenti

[S. Woolley, \*No, the human-robot singularity isn't here. But we must take action to govern AI\* | The Guardian, 10 febbraio 2026](#)

[T. Wheeler, \*Governing the AI transition: Lessons from the 1996 Telecommunications Act\* | Brookings, 9 febbraio 2026](#)

[A. Weber Pierson, \*Commentary: The Biases We Don't See — Preventing AI-Driven Inequality in Health Care\* | The Observer, 6 febbraio 2026](#)

[I. Mej, \*Model collapse, ovvero quando l'AI perde contatto con la realtà: il rischio è soprattutto psicologico\* | Il Fatto Quotidiano, 5 febbraio 2026](#)

[V. Rao, \*l'India e l'intelligenza artificiale\* | Corriere della Sera, 5 febbraio 2026](#)

[P. Benanti, \*Moltbook, la piattaforma dove pacatamente dialogano le intelligenze artificiali\* | Il Sole 24 Ore, 4 febbraio 2026](#)

[A. Puliafito, \*I chatbot affettivi sono davvero un problema?\* | Internazionale, 4 febbraio 2026](#)

[L. Pierin, \*L'intelligenza artificiale come progetto umano. Intervista a Luciano Floridi\* | Scienza in Rete, 2 febbraio 2026](#)

[J. Fo, \*L'AI non è cattiva: l'hanno programmata per fare soldi. Che importa se qualcuno poi si fa del male!\* | Il Fatto Quotidiano, 2 febbraio 2026](#)

[A. Dina, \*I data center AI e la rivolta silenziosa delle comunità: il vero collo di bottiglia dell'intelligenza artificiale\* | Rivista AI, 1 febbraio 2026](#)

[\*Il lavoro di chi rivede gli articoli scientifici si è molto complicato\* | Il Post, 28 gennaio 2026](#)

## Corsi, convegni e pubblicazioni

### Corsi

[Università Cattolica del Sacro Cuore, \*Dirigere con l'intelligenza artificiale: strumenti, strategie e rischi\* | Residenza Gasparini- Piacenza, 26-28 febbraio 2026](#)

[ISPI, \*L'Intelligenza Artificiale applicata alla Cooperazione internazionale\* | Web-live, 20- 21 marzo 2026](#)

### Eventi e convegni

[\*Presentazione del Rapporto Lombardia 2025 "Intelligenza Artificiale per la sostenibilità"\* | Auditorium Polis-Lombardia, 24 febbraio 2026](#)

### Pubblicazioni

[D. Boriati, M. D'Ambrosio, \*L'Intelligenza Artificiale tra scienza, umanità e società\* | FrancoAngeli, febbraio 2026](#)

[F. Petit, \*The future employment impact of artificial intelligence and emerging digital technologies in Europe\* | EU Publications, gennaio 2026](#)

Link attivi al 13 febbraio 2026

Prodotto da: PoliS-Lombardia

Coordinamento editoriale a cura di **Davide Perillo**

Comitato Scientifico: **Marco Sica, Marco Bassini, Annalisa Negrelli**

*(hanno collaborato: Beatrice Capitanio, Annaclara De Tuglie, Chiara Rizzo, Vanna Toninelli)*