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

PoliS Lombardia | Regione Lombardia

PoliS AI NEWS

Newsletter sull'Intelligenza Artificiale
a cura di PoliS-Lombardia

Anno III – n. 8/2026

In questo numero
In evidenza
Focus
Normativa
Applicazioni alla Pubblica Amministrazione
AI in pillole
Notizie
Commenti
Corsi, convegni e pubblicazioni

In questo numero

A che punto siamo nell'adozione dell'AI? Ne parliamo nella sezione "In Evidenza" grazie a quella che è considerata la "Bibbia" del settore, **l'Artificial Intelligence Index report della Stanford University**. Segue un approfondimento delle **strategie di implementazione dell'AI nel Parlamento italiano**, un'indagine sull'**affidabilità degli *AI Overviews* di Google** e un'analisi **dell'inaspettato rapporto tra la Gen-Z statunitense e l'Intelligenza artificiale**. Nel "Focus" un recap delle **posizioni di Consiglio, Parlamento e Commissione europei sul Digital Omnibus sull'AI** e i pronostici sull'esito finale. In "Normativa", la neo-adottata **legge regionale sarda sull'AI**. La "Pillola" entra nel merito delle **abilità di resilienza e robustezza che permettono ai sistemi AI di autocorreggersi**. Poi, come sempre, esempi, applicazioni alla PA, notizie e commenti. Buona lettura!

In evidenza

L'AI corre troppo, difficile tenere il passo (parola dello Stanford report…)

L'intelligenza artificiale corre sempre più veloce. *The 2026 Artificial Intelligence Index Report*, il report della Stanford University considerato per molti versi la "Bibbia" del settore, ogni anno fotografa lo sviluppo dei modelli a livello globale, non solo dal punto di vista tecnologico ma anche nei risvolti sociali, competitivi e di governance.

La nona edizione della ricerca analizza le novità del 2025 e stabilisce che **l'AI è il maggiore fattore di cambiamento di questo secolo**. L'AI generativa **ha raggiunto una diffusione del 53% in tre anni**, un ritmo esponenziale rispetto a quello che ha portato all'affermazione di pc e Internet.

Sul piano delle performance su **OSWorld, piattaforma creata da ricercatori dell'Università di Hong Kong che testa i modelli sottoponendo compiti informatici reali, registra una percentuale di successo passata dal 12 al 66%.** E c'è anche chi, come Gemini Deep Think, si aggiudica una medaglia d'oro alle Olimpiadi internazionali di matematica. In generale, **la capacità di calcolo globale nel settore dell'AI è cresciuta di 3,3 volte all'anno dal 2022, grazie ai nuovi data center.**

Tutto bene? Non proprio. Intanto piattaforme più strutturate di OSWorld registrano un terzo di fallimenti. Molto più banalmente, **se chiediamo a un modello di leggere un orologio analogico, sbaglierà in un caso su due** e le faccende domestiche nella realtà vengono completate solo nel 12% dei casi reali, contro l'89,4% dei test in laboratorio.

Ma è sugli aspetti più seri, come la sicurezza, che emergono ombre. **Gli incidenti legati all'AI sono passati dai 233 del 2024 ai 362 del report 2025** (AI Incident Database). Un test ha verificato che le allucinazioni – ossia l'incapacità di distinguere tra fatti e credenze – non sono episodiche, ma possono arrivare a verificarsi nel 94% dei casi. Inoltre, i tentativi di privilegiare un'Intelligenza artificiale responsabile – in grado di garantire sicurezza, equità, trasparenza – si scontrano con l'evidenza sperimentale che migliorando uno dei parametri, altri possono peggiorare.
Anche l'impatto ambientale nel 2025 è cresciuto: 72.816 tonnellate di CO_2 equivalente sono le emissioni generate da Grok 4, **il consumo idrico per GPT-4o supererebbe il fabbisogno di acqua potabile di 12 milioni di persone.** I data center hanno una potenza di 29.6 GW, pari a quanto consuma lo stato di New York nei momenti di picco.

Lo sviluppo dell'AI, naturalmente, ha anche un impatto geopolitico. **La competizione per il primato nel settore coinvolge Stati Uniti e Cina, che si alternano sul gradino più alto del podio**: gli Usa vantano una produzione più corposa e di maggiore impatto, mentre la Cina è in testa per numero di pubblicazioni, citazioni, produzione di brevetti e installazione di robot industriali.

Ma come sta andando la capacità di governare uno sviluppo così rapido? **Gli standard per valutare l'AI progrediscono meno rispetto all'Intelligenza artificiale stessa, diventando obsoleti nel giro di sei mesi**. Studi

recenti fanno affiorare dubbi anche sulla affidabilità delle piattaforme, per la capacità dei modelli di adattarsi ai sistemi di verifica e di riuscire a "ingannarli". Humanity's Last Exam, benchmark studiato per "mettere in difficoltà" l'AI con domande particolarmente complesse, ha registrato nei modelli esaminati un miglioramento di 30 punti percentuali. Ciò ha consentito di arrivare a un'accuratezza del 38%, un risultato ancora deludente in valori assoluti.

Le nazioni in cui cresce più rapidamente l'introduzione di politiche sull'AI sono quelle che finora erano meno strutturate in tal senso. Due velocità che ritroviamo anche nella capacità di assicurarsi il controllo delle infrastrutture che alimentano l'Intelligenza artificiale. **L'Europa appare prudente in entrambi i casi: policy molto articolate e investimenti inferiori rispetto alla potenza di fuoco messa in campo da Stati Uniti, Cina e non solo**.

Gli Usa hanno 5.427 data center, dieci volte di più degli altri Paesi, e primeggiano anche negli investimenti privati, con 285,9 miliardi di dollari nel 2025 (+127,5% e +200% per l'AI generativa) e 1.953 nuove aziende d'Intelligenza artificiale. La **Cina** si ferma a 12,4 miliardi, ma in questo caso le risorse pubbliche hanno un ruolo importante, considerando che **i fondi governativi hanno investito circa 184 miliardi di dollari** tra il 2000 e il 2023.

In materia d'investimenti pubblici gli Stati Uniti tra il 2013 e il 2024 hanno investito circa 20,4 miliardi di dollari. **L'Europa, dal canto suo, ha sottoscritto contratti tra il 2013 e il 2024 per 3,7 miliardi di dollari.** Quasi la metà (1,6 miliardi) provengono dal Regno Unito, seguito da Germania (505 milioni) e Francia (320 milioni).

Il panorama fa pensare che sarà sempre più difficile vedere alleanze internazionali, perché **l'AI e le sue infrastrutture sono sempre più una leva di potere che invita al sovranismo**. Questo limita la possibilità di governare il fenomeno dal punto di vista delle istituzioni, e non solo dei privati. E di fronte a uno strumento tanto potente, non è un fattore trascurabile.

*Il testo del rapporto
Per approfondire:
L. Tremolada, *L'Ai Index Report 2026 di Stanford in cinque grafici* |Il Sole 24 Ore, 19 aprile 2026

AI e politica: il Parlamento alla prova dell'algoritmo

a cura di Annalisa Negrelli

L'utilizzo dell'AI nelle istituzioni e nei processi democratici cresce in modo sempre più pervasivo e interessa anche le istituzioni parlamentari, **con un impatto significativo sulla produzione normativa, la gestione dei dati e la qualità del controllo democratico**. Ma a che punto siamo in Italia? Per capirlo, è utile tornate al dossier presentato alcuni mesi fa a Bruxelles dal Parlamento italiano, che fa il punto della situazione anche in una prospettiva comparata.

Lo studio evidenzia come la cornice regolamentare Ue imponga ai **parlamenti degli Stati membri**, che scelgono di adottare strumenti di AI, di **adeguare i processi interni ai nuovi standard di sicurezza, tracciabilità e supervisione umana**; di coordinarsi con l'AI Office europeo e di definire in modo più puntale le procedure interne per l'uso di strumenti generativi e analitici.

Gli strumenti di AI adottati negli altri parlamenti Ue sono molto eterogenei come emerge dal Dossier comparato sull'utilizzo AI in Senato (in Austria, EULE media monitor; HANS, in Estonia; in Francia LLaMandement; **ChatGPT** in Slovenia), incentrati principalmente sulla **trascrizione automatica delle sedute** tramite sistemi ASR, che velocizzano la produzione dei resoconti (il più evoluto è Archibot, usato dal Parlamento europeo). Nessun parlamento ancora utilizza software AI per la **valutazione dell'impatto normativo**, il **monitoraggio dell'attuazione delle leggi** o la **gestione automatica dell'agenda**, soprattutto per ragioni di privacy e sicurezza.

Per quanto riguarda **l'esperienza italiana**, dal dossier emerge che **il Senato ha già avviato da tempo un percorso strutturato di introduzione dell'AI**, con un **approccio orientato a governance, mitigazione del rischio e qualità**

dei dati.
Le principali linee operative si articolano in alcune direzioni essenziali: analisi normativa assistita (strumenti per la comparazione legislativa, per l'estrazione di riferimenti, per l'individuazione di impatti e correlazioni); supporto documentale (generazione di sintesi, briefing e materiali per Commissioni e Aula); ottimizzazione dei flussi informativi (classificazione automatica degli atti, miglioramento dell'accessibilità alle banche dati); definizione di linee guida interne (con focus su principio di trasparenza, supervisione umana, tracciabilità delle fonti, divieto di delega decisionale all'algoritmo).

Se il Senato sembra orientato verso un modello di AI inteso come strumento di supporto tecnico – ponendo grande attenzione alla responsabilità amministrativa e alla coerenza con i principi dell'AI Act e privilegiando un approccio più prudenziale e orientato alla sperimentazione controllata –, la Camera dei Deputati punta su un ecosistema digitale integrato e su una maggiore automazione dei flussi documentali. In questo caso, le aree di intervento includono ambiti quali: AI generativa per la redazione di documenti (strumenti per note, sintesi, schede tecniche es. NORMA, assistente virtuale per l'analisi della produzione legislativa); classificazione automatica degli atti (sistemi di tagging e smistamento basati su modelli linguistici); progetti di data governance (standardizzazione dei dataset, qualità dei metadati, sicurezza dei flussi informativi); formazione del personale (sviluppo di competenze interne per un uso consapevole e conforme dell'AI); strumenti di semplificazione (MSE, sistema di scrittura assistita per emendamenti parlamentari; DEPUCHAT, chatbot per facilitare l'accesso dei cittadini alle attività dei deputati). Oltre che l'elaborazione di un codice di condotta specifico, improntato alla trasparenza verso i cittadini e alla prevenzione dei bias algoritmici.

Entrambi i rami del Parlamento, quindi, ricorrono all'AI come supporto tecnico e non come strumento sostitutivo della funzione politica, come dimostrano l'attenzione alla supervisione umana e alla tracciabilità, ma anche lo sviluppo di linee guida interne coerenti con l'AI Act e, infine, la centralità della qualità dei dati e della sicurezza informatica.

Per approfondire:
eTools for regulatory simplification and consistency | European Parliamentary Research Service, Marzo 2026
La Camera presenta i primi tre prototipi di Intelligenza Artificiale per il Parlamento. Lectio magistralis del professor Floridi | Camera dei deputati, Luglio 2025
Ufficio valutazione impatto, *L'intelligenza artificiale in Senato* | Senato della Repubblica

Google ci risponde, ma dice la verità? Dentro i limiti degli AI Overviews

Quando cerchiamo qualcosa su Google, sempre più spesso la prima cosa che vediamo è un riquadro di testo: un riassunto generato dall'AI, i cosiddetti *AI Overviews*. Ma quanto sono affidabili queste risposte automatiche? A provare a misurarlo è stata un'inchiesta del New York Times, condotta insieme alla startup Oumi, una piattaforma open source di Intelligenza artificiale fondata da ex ingegneri e ricercatori di Google e Apple.

L'analisi mostra che gli AI Overviews risultano corretti in circa il 90% dei casi. Un dato che, isolato, potrebbe sembrare rassicurante. Il problema emerge però se si guarda alla scala: Google gestisce circa 5000 miliardi di ricerche all'anno. Anche un margine di errore "solo" del 10% si traduce così in circa 57 milioni di risposte sbagliate ogni ora, con un potenziale enorme di diffusione di informazioni imprecise o fuorvianti.

Per condurre lo studio, Oumi ha utilizzato *SimpleQA*, un test di benchmark progettato da OpenAI e ampiamente usato nel settore per valutare l'accuratezza dei modelli di Intelligenza artificiale. I test sono stati effettuati in due momenti: a ottobre, quando gli AI Overviews erano basati sul modello Gemini 2, e a febbraio, dopo il passaggio al più avanzato Gemini 3. In entrambi i casi sono state analizzate 4.326 ricerche.

I risultati mostrano un miglioramento: l'accuratezza passa dall'85% con Gemini 2 al 91% con Gemini 3. Google ha però criticato l'analisi, definendola poco rappresentativa delle reali ricerche degli utenti e basata su un benchmark che conterrebbe errori. Eppure, come rileva l'inchiesta, la stessa Google ha pubblicato test interni su Gemini 3 che indicano un tasso di informazioni errate fino al 28% – una percentuale ben più alta di quella stimata da Oumi.

C'è poi un altro aspetto forse ancora più problematico. Secondo l'analisi, **con Gemini 3 aumentano le risposte formalmente corrette ma "non fondate"**: informazioni giuste nei fatti, ma che rimandano a fonti che in realtà non le supportano. Questo **fenomeno riguarda il 56% dei casi**, contro il 37% registrato con Gemini 2. Percentuali tutt'altro che marginali. Inoltre, **tra le oltre 5.000 citazioni analizzate** durante l'indagine, **Facebook e Reddit emergono come la seconda e la quarta** fonte più utilizzate dagli AI Overviews.

Sono dati utili, certo. Ma le criticità non sono solo tecniche. Come ha raccontato The Guardian già nell'ottobre 2025, in vari Paesi europei gli stessi editori hanno lanciato l'allarme: **gli AI Overviews riducono drasticamente i clic verso i siti di informazione**. Secondo uno studio della società di analisi Authoritas, l'introduzione di questi riepiloghi ha comportato quasi il 50% di clic in meno su desktop e il 37% in meno su mobile per il traffico web degli editori di notizie nel Regno Unito. Un'analisi del Pew Research Center conferma il calo, mostrando che gli utenti cliccano su un link all'interno di un riepilogo generato dall'AI solo una volta ogni cento.

Non sorprende quindi che, nello stesso periodo, la Federazione Italiana Editori Giornali (FIEG) abbia definito Google un vero e proprio "traffico-killer" e presentato un **reclamo formale all'Agcom** contro il servizio AI Overviews, che si pone non solo «in diretta concorrenza con i contenuti prodotti dalle aziende editoriali», ma provoca anche «una riduzione della loro visibilità e reperibilità».

Più esposti all'AI, ma meno convinti: i giovani americani fanno marcia indietro?

La Gen-Z statunitense è sempre più scettica sui benefici dell'AI: non solo **l'adozione è in una fase di stallo** – ad usarla **almeno una volta alla settimana è il 51%,** dato pressoché identico a quello del 2025 –, ma **calano entusiasmo** (-14% rispetto all'anno scorso) e **ottimismo** (-9%), mentre **cresce un sentimento di rabbia** (+9%).

È quanto emerge dal report *The AI Paradox: More Exposure, Less Confidence Among Gen Z*, condotto da Gallup sugli **atteggiamenti di più di 1.500 giovani americani tra i 14 e i 29 anni nei confronti dell'Intelligenza artificiale.** «**La Generazione Z è sempre più scettica e negativa**, pur partendo da una situazione in cui già l'anno scorso non era particolarmente ottimista», ha affermato Zach Hrynowski, ricercatore presso Gallup, in un'intervista al New York Times.

Le ragioni di questo scetticismo sono diverse. Da un lato, sono sempre di più i ragazzi convinti che **l'AI faccia più male che bene alla creatività (38%) e al pensiero critico (42%)**; dall'altro, invece, sono sempre **meno quelli persuasi che possa aumentare l'efficienza nell'apprendimento** (-10 punti percentuali rispetto al 2025) e anzi, **l'80% teme che possa in qualche modo renderlo più difficile** a causa della perdita di competenze che vengono esternalizzate all'AI.

Chi frequenta le *high schools,* però, è anche convinto che **all'università sarà fondamentale saper usare l'AI (52%)** e crede che quando il momento arriverà, sarà in grado di farlo (56%). Ma da dove arriva tutta questa sicurezza? Sono **sempre di più le scuole che adottano policy sull'uso dell'Intelligenza artificiale** (dal 51% nel 2025 al 74% nel 2026), offrendo la possibilità di usarla direttamente sui computer scolastici (49%) o per fare i compiti (65%).

Sul lavoro le cose non migliorano: **quasi la metà dei giovani (48%) ritiene che i rischi superino i benefici**. Solo il 15% pensa il contrario e il restante 37% crede che si equivalgano. Il motivo? **Ci si fida sempre meno degli output prodotti dall'AI** o con il suo supporto: la maggior parte (il 69%) ritiene di gran lunga più attendibile l'attività svolta solamente da colleghi umani.

Ma non è finita qui: dalle interviste di alcuni partecipanti allo studio, riportate in un articolo del New York Times, emerge anche una **preoccupazione per l'accesso alle posizioni lavorative iniziali**, quelle a cui i giovani possono puntare ma che **spesso** sono **rimpiazzate dall'AI**. «Ho l'impressione che qualsiasi cosa mi interessi possa essere sostituita, magari già nei prossimi anni», ha spiegato Sydney Gill, una matricola alla Rice University di Houston, ricordando il suo ottimismo iniziale quando al liceo guardava all'AI come a uno

strumento di apprendimento. Ora che deve scegliere un percorso di studi in funzione del mercato del lavoro, non è più sicura del "bene" dell'AI e il suo entusiasmo inizia a traballare…

*il testo del report *The AI Paradox*

Focus

Il Digital Omnibus verso il trilogo: le posizioni emerse e i pronostici sull'esito finale

Il Digital Omnibus sull'AI è entrato nella sua fase più propriamente negoziale. La **proposta della Commissione**, presentata il 19 novembre 2025, **mirava** soprattutto **a rendere più "digeribile" l'attuazione dell'AI Act senza alterarne l'impianto** di fondo: **proroga mirata dei termini** di applicazione per i sistemi AI ad alto rischio, estensione di alcune **misure di alleggerimento alle piccole e medie imprese**, disciplina di dettaglio per il **trattamento di categorie particolari di dati personali** per finalità di identificazione e **correzione di distorsioni, rafforzamento dei poteri dell'AI Office** e altri correttivi volti complessivamente a ridurre frammentazione e oneri per le imprese. La logica di fondo della Commissione è quindi quella della semplificazione e della certezza del diritto, in linea con l'agenda europea sulla competitività nel digitale.

Su queste basi, il **Consiglio e il Parlamento hanno però già impresso due direzioni parzialmente diverse**. Il **Consiglio**, pur mantenendo le linee essenziali della proposta della Commissione, ha adottato **un approccio di maggior rigore**. In particolare, ha trasformato il **rinvio dei termini** per i sistemi ad alto rischio in una scansione temporale fissa, fissando il **2 dicembre 2027** per i sistemi ad alto rischio "standalone" e il **2 agosto 2028** per quelli incorporati in prodotti. Ha inoltre **reintrodotto l'obbligo di registrazione nella banca dati Ue** anche per i fornitori che ritengano il proprio sistema escluso dalla qualifica di sistemi ad alto rischio, ha ripristinato il **criterio della stretta necessità per il trattamento delle categorie particolari di dati ai fini della correzione di distorsioni** e, da ultimo, ha rinviato al 2 dicembre 2027 l'istituzione delle sandbox nazionali, oltre a chiedere alla Commissione specifiche linee guida per gli operatori soggetti anche alla legislazione settoriale armonizzata. Infine, ha **inserito un nuovo divieto relativo ai sistemi che generano contenuti a contenuto sessuale** o intimi non consensuali e materiale di abuso sessuale su minori, sulla scorta forse degli echi delle vicende del chatbot Grok.

Il **Parlamento europeo**, dal canto suo, si è mosso in una **direzione in parte più politica e più chirurgica**. Emerge anzitutto il **sostegno a un nuovo divieto delle applicazioni di "*nudification*"**, cioè dei sistemi che creano o manipolano immagini sessualmente esplicite o intime di persone identificabili senza il loro consenso, salvo il

caso in cui siano previste misure di sicurezza efficaci per impedirne alla fonte la generazione. Sostiene poi una **maggiore flessibilità per le imprese**, compresa l'estensione di varie misure di supporto alle piccole e medie imprese, e accoglie l'idea di **consentire il trattamento di dati personali per individuare e correggere distorsioni** nei sistemi AI, ma insistendo sul fatto che ciò debba avvenire **solo quando strettamente necessario** a tale scopo. Inoltre, rispetto all'assetto originario della Commissione, il Parlamento sembra voler enfatizzare la **necessità di evitare sovrapposizioni tra l'AI Act e la normativa settoriale di sicurezza** dei prodotti, attenuando gli obblighi nei casi in cui il prodotto sia già fortemente regolato.
Il Parlamento ha provato anche a **rafforzare l'apparato di accompagnamento sull'alfabetizzazione** all'AI, ipotizzando linee guida operative, sostegno specifico alle PMI e un ruolo più marcato del Board nell'adozione di raccomandazioni comuni. Su questo punto, il Consiglio appare, invece, più misurato e più vicino alla linea di fondo proposta dalla Commissione.

In sintesi, il confronto tra i tre testi mostra una **convergenza** sul fatto che il **Digital Omnibus non debba riscrivere l'AI Act, ma preparane e favorirne l'attuazione**; divergono però i contenuti di alcune modifiche. La Commissione resta ancorata a una semplificazione "tecnica"; il Consiglio privilegia una migliore scansione temporale degli obblighi e un riordino delle competenze; il Parlamento intende introdurre segnali più visibili su aspetti sensibili, in particolare sul fronte della *nudification* e del sostegno alle imprese. Il Consiglio ha dichiarato che la Presidenza avvierà i negoziati con il Parlamento, mentre il Parlamento ha affermato di essere pronto a iniziare i negoziati con il Consiglio sulla forma finale del testo. La **fase successiva è dunque quella del trilogo, da cui si verificherà il grado di adesione alla semplificazione originariamente proposta dalla Commissione** o la prevalenza di un approccio più interventista.

Normativa

La Sardegna approva la Legge regionale sull'AI

Approvata dalla Regione Sardegna il 12 marzo 2026, **la legge regionale n. 6 del 2026 disciplina lo sviluppo, la ricerca e l'utilizzo dell'Intelligenza artificiale in ambito regionale**, promuovendo l'uso etico, antropocentrico e sostenibile. Una tra le prime regioni a dotarsi di una regolamentazione specifica in tema di AI, la Sardegna ambisce a diventare polo di riferimento sul territorio italiano con la costituzione, tra l'altro, di un **Hub AI regionale dedicato (HUBIAS) e di Hub territoriali,** in collaborazione con il CRS4, Sardegna Ricerche e il sistema delle Università regionali, per ricerca, innovazione e supporto alle imprese.
La legge individua **quattro principali direttrici di intervento**:

- sostegno al progresso socioeconomico attraverso il rafforzamento delle competenze digitali e il supporto alle imprese;
- miglioramento della qualità dei servizi pubblici, favorendo soluzioni innovative che rendano la pubblica amministrazione più efficiente e accessibile, **supportando tra l'altro anche la diagnostica medica**;
- adozione responsabile dell'AI nei settori strategici;
- introduzione di misure dedicate alla formazione e alla riqualificazione professionale.

Per approfondire
Il testo della legge n. 6 del 2026
Sardegna, una legge sulla IA nella pubblica amministrazione e nella società | Ansa, 10 marzo 2026

Applicazioni alla Pubblica Amministrazione

ITALIA

FVG – progetto ChatBot edih-PAI

Progetto sperimentale di Chatbot AI conversazionali della Regione Friuli-Venezia Giulia, avviato il 28 novembre 2025, per semplificare l'accesso ai servizi comunali per le pratiche amministrative, anagrafiche, documentali e urbanistiche, operativo in 9 Comuni. L'obiettivo è semplificare il rapporto tra cittadini e amministrazioni, fornendo risposte 24 ore su 24 e in più lingue.

edih-PAI

Chatbot intelligenti su 9 siti web comunali: Insiel avvia la sperimentazione | Anci FVG, 25 novembre 2025

UNIONE EUROPEA

Børge – Danimarca

Assistente alla scrittura basato sull'AI della Pubblica amministrazione danese, è utilizzato da circa 40 autorità pubbliche per ottimizzare e rendere più chiari i contenuti dei portali nazionali.

Børge - AI-skriveassistent

MONDO

FAIr LAC – Uruguay

Il progetto mira a promuovere e rafforzare l'uso responsabile dell'Intelligenza artificiale, in particolare nell'amministrazione pubblica, attraverso l'attuazione della Strategia di AI per il Governo Digitale.

Uruguay | fAIrLAC

AI in pillole

Resilienza e robustezza: le abilità dell'AI di correggersi e ritararsi

a cura di Annalisa Negrelli

La resilienza digitale (o capacità di recupero) indica **l'abilità di un sistema AI di recuperare rapidamente uno stato operativo adeguato** dopo un incidente tecnico, di sicurezza o operativo, come un attacco informatico, un guasto o un imprevisto di varia natura. A differenza della semplice *resistenza*, un sistema AI *resiliente* **accetta la possibilità del fallimento o dell'intrusione** (in situazioni ad esempio di fallback, ridondanze, graceful degradation) e si concentra sulla **continuità operativa** e su **procedure di rapido ripristino che minimizzino impatti e inefficienze su utenti e terzi.** Gli standard europei richiedono che i sistemi ad alto rischio siano resilienti a errori interni, incongruenze e interazioni con persone o altri sistemi e che, al fine di ridurre il rischio, adottino misure tecniche o organizzative adeguate – anche rispetto a tentativi di terzi di manipolarne comportamento e output. La resilienza di un modello AI si testa mettendo alla prova il sistema, con scenari avversi, esercitazioni e test periodici.

Al pari della resilienza AI, **la robustezza tecnica** (o prevenzione del danno locale) **è anch'essa un pilastro fondamentale nei framework di governance dell'AI** e consiste nella **capacità di resistere all'urto e di mantenere il livello di prestazioni "sotto ogni circostanza" entro i limiti dichiarati**. Una robustezza adeguata include resistenza a situazioni peculiari (rumore, drift dei dati), condizioni specifiche e attacchi informatici mirati. Anche in questo ambito, gli standard Ue richiedono che i sistemi ad alto rischio siano robusti e che, a tal fine, adottino sistemi di "ridondanza" per reagire a situazioni critiche, quali procedure di back up o fail safe. Questa capacità del sistema, proprio come la resilienza, si misura con metriche, prove e stress-test di sicurezza.

Casi concreti di essenzialità della resilienza e robustezza AI

Nel **settore energetico**, ad esempio, la **resilienza digitale** non è tanto un obiettivo tecnico, quanto piuttosto una **necessità concreta** perché le centrali automatizzate si basano generalmente su apparati complessi (sistemi di controllo industriale, infrastrutture IT di gestione e algoritmi di machine learning). Ciò significa che un attacco informatico a un singolo segmento di controllo può generare blackout o anomalie di rete con effetti sistemici che impattano sull'intero sistema.
Anche il settore della **sanità digitale** rappresenta un fronte particolarmente sensibile. Le apparecchiature medicali più complesse dialogano con sistemi IT ospedalieri e moduli di AI per diagnosi automatizzate: un'eventuale interruzione o guasto può influire concretamente sulla sicurezza dei pazienti.

E, ancora, **città intelligenti**, **catene di approvvigionamento** nel settore di logistica e trasporto, **piattaforme di gestione del traffico**, della **nettezza urbana**, della **sicurezza**, sono tutti frangenti che evidenziano in concreto l'essenzialità della resilienza AI, come emerge dai più recenti studi in proposito. D'altro canto, anche la robustezza AI può essere particolarmente critica in applicazioni di vitale importanza **come i sistemi di combattimento, i droni di ricerca e soccorso o il triage sanitario supportato dall'AI.**

Robustezza e resilienza sono dunque caratteristiche essenziali e complementari: **un sistema AI robusto deve anche essere cyber-resiliente** e deve essere progettato per degradare in modo sicuro, con allarmi tempestivi e meccanismi di blocco controllato, come stabilito anche dalle normative Ue che hanno reso obbligatoria la dimostrazione della capacità di ripristino e continuità dei modelli AI in ambienti sempre più complessi.

Per approfondire:
A. Bertolino, *Privacy, sicurezza, resilienza: le parole d'ordine per l'IA in Europa* | Agenda Digitale, 3 ottobre 2025
A. Badman e M. Kosinski, *Che cos'è la resilienza delle applicazioni?* | IBM
Articolo 15: Accuratezza, robustezza e cibersicurezza | AI Act Service Desk
A. McGrath e A. Jonker, *Che cos'è la sicurezza dell'AI?* | IBM
N. Panigada, *Principi generali dell'AI: accuratezza, robustezza e cibersicurezza. Commento all'.art. 15* | Commentario tematico del regolamento (Ue) 2024/1689 e della legge italiana 23 settembre 2025, n. 132, 2026

Notizie

G. Riva, *SpaceX è pronta ad acquistare Cursor per 60 miliardi di dollari* | Il Sole 24 Ore, 22 aprile 2026

L. Tremolada, *Quali sono le competenze per lavorare con l'AI?* | Il Sole 24 Ore, 21 aprile 2026

T. Brambilla, *Anthropic-amministrazione USA: dialogo riaperto in nome della cybersecurity* | AI News, 20 aprile 2026

A. Longo, *Le start-up acquisite dalla Big Tech svelano il futuro dell'intelligenza artificiale* | Il Sole 24 Ore, 19 aprile 2026

F. Cella, *«Se supporti Trump, ti sei appena fatto un amico»: la propaganda invade i social con finti americani creati con l'Ai* | Corriere della Sera, 18 aprile 2026

I. Miralli, *I chatbot con intelligenza artificiale hanno effetti sugli adolescenti simili alle droghe* | Corriere della Sera, 17 aprile 2026

D. Zorzetto, *I farmaci? Li prescrive l'Intelligenza artificiale* | La Repubblica, 17 aprile 2026

P. Raimondi, *ASML accelera, tra chip, AI e geopolitica* | Rivista AI, 16 aprile 2026

M. Rovelli, *«Il riconoscimento facciale sui Ray-Ban Meta è un'arma per stalker e predatori sessuali»: l'allarme di 77 organizzazioni civili* | Corriere della Sera, 15 aprile 2026

G. Clericò, *AI e lavoro: Google lancia "AI Works for Italy", 2 milioni per formare tredicimila universitari italiani* | AI4Business, 15 aprile 2026

Humanitas, l'AI aiuta i medici durante le visite: così l'assistente virtuale trascrive il referto | Il Giorno, 13 aprile 2026

F. González, *Bixonimania, i chatbot si sono "bevuti" una malattia inventata di sana pianta (che poi è finita nei paper scientifici)* | Wired, 12 aprile 2026

P. L. Pisa, *Molotov contro l'abitazione di Sam Altman: arrestato un ventenne* | La Repubblica, 11 aprile 2026

Commenti

C. M. Medaglia, *Neurodivergenze e IA: chi sta imitando chi?* | La Repubblica, 22 aprile 2026

M. Burgess, *Nudi deepfake nelle scuole, la piaga è diventata una crisi globale* | Wired, 21 aprile 2026

L. Mei, *Fei-Fei Li e l'intelligenza artificiale centrata sull'umano: una filosofia che mette l'uomo al centro della macchina* | Rivista AI, 19 aprile 2026

F. Cella, *Come salvare le lingue invisibili nell'era dell'AI* | Corriere della Sera, 18 aprile 2026

H. Lonas, *We finally built the time machine. Now we're wasting the time we created* | CIO, 17 aprile 2026

L. Elliott, *AI is destroying jobs – and the energy crisis could make that much worse* | The Guardian, 16 aprile 2026

A. Puliafito, *Ha senso una costituzione delle intelligenze artificiali?* | Internazionale, 15 aprile 2026

P. Benanti, *IA e cyber risk: la fine dei modelli tradizionali di gestione del rischio* | Il Sole 24 Ore, 15 aprile 2025

L'ia che può difendere o violare tutti i software del mondo | Internazionale, 14 aprile 2026

A. Preiti, *È l'algoritmo non il social che ci rende schiavi* | Corriere della Sera, 14 aprile 2026

P. Benanti e F. Maffettone, *AI, il conflitto nascosto è sulle risorse* | Corriere della Sera, 13 aprile 2026

B. Agranovsky, *Why honest students fear AI detectors* | The Washington Post, 13 aprile 2026

O. Pollicino, *Cosa vuol dire eguaglianza oggi, il vero significato dell'articolo 3 della Costituzione al tempo degli algoritmi* | Wired, 11 aprile 2026

Corsi, convegni e pubblicazioni

Corsi

Promo PA, *Intelligenza artificiale e Pubblica Amministrazione: tecnologia, EU AI Act, L. 132/2025, prompt engineering - IX Edizione* | Online, 8-15-19-22-28 maggio 2026

Ca' Foscari Challenge School, *Agenti di Intelligenza artificiale per la PA: progettare e governare il lavoro digitale* | Online, 8 giugno- 15 giugno 2026

Eventi e convegni

ANCI Lombardia e SmartCityLab Milano, *Innovazione urbana e Intelligenza Artificiale Generativa* | Milano, 6 maggio 2026

Polimi School of Management, *L'Intelligenza Artificiale nella pubblica amministrazione* | Online, 7 maggio 2026

AI Week | Milano- Fiera Rho, 19-20 Maggio 2026

Pubblicazioni

L. Mari, S. Sancassiani, D. Bellasio, *L'intelligenza artificiale di Pinocchio. Ha ancora senso apprendere nell'era dei chatbot?* | Il Sole 24 Ore, aprile 2026

A. Chella, *Può un robot emozionarsi? Riflessioni sull'intelligenza artificiale, la coscienza e le emozioni nelle macchine* | Mondadori Università, aprile 2026

Link attivi al 24 aprile 2025

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)