



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 II – n. 16/2025

In questo numero

In evidenza

Focus

Normativa

Applicazioni alla Pubblica Amministrazione

AI in pillole

Notizie

Commenti

Corsi, convegni e pubblicazioni

In questo numero

Ma l'Italia è in grado di **attirare talenti nel campo dell'Intelligenza artificiale**? Stando ad una ricerca di Interface, pare di sì (anche se con qualche problema). Ne parliamo "In evidenza", accanto a un intervento sul ruolo

previsto per le autorità di garanzia nella **nuova legge italiana sull'AI** (mentre in "Normativa" se ne approfondisce un punto specifico: l'impatto sulla compliance nella PA) ad alcuni approfondimenti **sull'uso (e il possibile abuso) dell'AI in sanità**. Il "Focus" è su una **sentenza del Tribunale di Torino che apre scenari sull'uso dell'AI nei procedimenti giudiziari**. La "Pillola didattica" spiega i **segreti delle "scatole nere"** (e delle "scatole bianche") dei sistemi di Intelligenza artificiale. Poi, come al solito, notizie, commenti, segnalazioni... Buona lettura!

In evidenza

La "Dolce Vita" che attrae ma non trattiene... lo dicono i talenti AI

L'Italia (un po' a sorpresa) attrae parecchio i professionisti dell'AI, ma barriere economiche e problemi strutturali spingono molti a guardare altrove. È il **paradosso emerso dallo studio [La Dolce Vita Paradox](#)** del think tank europeo **Interface**, che ha analizzato **25.000 profili di talenti AI** e intervistato **esperti del settore** per indagare **lo stato del mercato italiano dell'Intelligenza artificiale**.

Un primo dato riguarda **la distribuzione dei talenti AI nel paese, ancora fortemente sbilanciata tra Nord e Sud**. La **Lombardia da sola conta oltre 7.000 specialisti**, mentre regioni come **Sicilia, Basilicata o Molise non arrivano a mille**. Nemmeno rapportando i dati alla popolazione il quadro cambia: Piemonte, Emilia-Romagna e Trentino-Alto Adige superano di gran lunga la media nazionale, confermando la concentrazione settentrionale. **A cosa si deve questo vantaggio? Università forti, più centri di ricerca e una maggiore densità di industrie tech** ([Microsoft, per esempio, vi ha concentrato i suoi progetti infrastrutturali](#)).

Milano guida la classifica delle città più attrattive. Il capoluogo lombardo è ormai riconosciuto come **uno dei principali hub europei dell'innovazione**, con **otto università**, oltre **232.000 studenti** – **il 7,4% internazionali**, **metà dei quali iscritti a corsi STEM** – e una **concentrazione record di professionisti AI**. Il numero di specialisti ICT in Lombardia è cresciuto del **9,7% rispetto al 2022** e del **24,3% rispetto al 2017**, portando il settore a rappresentare il **4,7% della forza lavoro regionale**. Dal **2019** la regione ha anche attirato la **quota maggiore di investimenti di venture capital** in Italia, consolidando **Milano come capitale nazionale delle startup**, le quali **tra gennaio e settembre 2022 hanno raccolto oltre 1,3 miliardi di euro in investimenti**, con un incremento del **90%** rispetto all'intero 2021, pur restando dietro giganti come Parigi e Londra.

I punti di forza sono evidenti, ma **non mancano anche le criticità**. L'**aumento del costo della vita** (senza un corrispondente incremento degli stipendi) e la **pressione immobiliare** rischiano di ridurre la capacità di attrattività. A questo si aggiungono, secondo il report, la **burocrazia complessa**, la **preferenza delle aziende per candidati che parlano italiano** e le **limitate prospettive di carriera**, fattori che frenano ulteriormente l'ingresso di talenti internazionali.

Attualmente **l'87,9% della forza lavoro AI in Italia è domestica**. Tra i **professionisti stranieri** del settore prevalgono quelli provenienti **da Argentina, Iran e Regno Unito**. In particolare, **l'1,6% arriva dall'Argentina**, a conferma della storica presenza di comunità italo-discendenti in America Latina.

E i talenti AI italiani? Il flusso in uscita è ancora significativo. **La Svizzera si conferma la destinazione preferita**, complice la vicinanza geografica, gli stipendi più alti e la lingua condivisa. Molti scelgono **anche Germania, Regno Unito e Francia**, dove i percorsi di carriera sono più rapidi e le retribuzioni decisamente superiori.

Un'ultima particolarità italiana riguarda la **presenza femminile**. Le donne **superano gli uomini nei risultati accademici** e nelle **lauree STEM avanzate**, e l'Italia è il Paese OECD con la **più alta percentuale di articoli**

scientifici AI firmati da almeno un'autrice. Milano si distingue in Europa con il **30,7% di professioniste AI**. Tuttavia, l'entusiasmo iniziale **non si traduce in carriere durature: la quota femminile scende dal 35,4% nelle posizioni entry-level al 14,2% in quelle senior**. La **precarietà contrattuale** colpisce più le donne e i **salari restano inferiori**: nei primi anni dopo la laurea i contratti a tempo indeterminato sono meno frequenti per loro, 49,9% contro il 56,1% degli uomini.

*Il [link](#) al rapporto

La legge italiana sull'AI: il perché della scelta di due autorità competenti

In questa clip a cura del prof. Marco Bassini dell'Università di Tilburg approfondiamo le scelte compiute dal legislatore italiano nella designazione delle autorità nazionali competenti in materia di intelligenza artificiale, cercando di cogliere le motivazioni di fondo per la scelta di Agid e ACN e l'equilibrio di competenze che verrà a disegnarci con le altre autorità indipendenti come il Garante privacy

*[Video Le autorità nella legge AI](#)

AI e medicina: così l'algoritmo fa lavorare meglio gli ospedali...

Una **riduzione del 24% del tempo** che i medici dedicano alla **stesura delle note cliniche** e del **17% del lavoro fuori orario**: questi sono i risultati ottenuti grazie all'introduzione di strumento di AI ambientale – che genera appunti clinici basandosi sulle conversazioni con i pazienti – presso l'azienda ospedaliera Northwestern Medicine di Chicago, Stati Uniti. È un traguardo importante visto che i **professionisti sanitari dedicano molto tempo**, forse troppo, **alla compilazione dei documenti e delle cartelle cliniche** (tra il 20% e il 60% del loro orario lavorativo).

Questa è solo una delle tante **applicazioni dell'AI in ambito sanitario mappate dalla Commissione Europea** nel report [Study on the deployment of AI in healthcare](#), che fa luce sui **possibili contributi che l'AI può dare per affrontare le sfide che le organizzazioni sanitarie devono affrontare**, tra cui l'invecchiamento della popolazione, la crescente diffusione di patologie croniche, l'aumento dei costi e la carenza del personale. Ma in che modo può tornare utile?

A livello generale, «**migliorando l'efficienza operativa, riducendo gli oneri amministrativi e potenziando i percorsi diagnostici e terapeutici**». Nel report poi vengono riportati tutta una serie di casi pratici in cui l'AI sta iniziando a fare la differenza.

Per quanto riguarda l'**incremento dell'efficienza**, un esempio significativo si trova negli **Stati Uniti al John Hopkins University Hospital**. Qui l'AI analizza le cartelle cliniche, facilitando il recupero delle informazioni sui pazienti e pianificando gli appuntamenti. Questo ha permesso una riduzione dei tempi di assegnazione dei letti al pronto soccorso del 30%, dei ritardi nei trasferimenti in sala operatoria del 70% e, infine, degli intervalli di risposta delle ambulanze di 63 minuti.

Sempre in questo senso, **l'AI può fornire assistenza nella fase di triage**, stabilendo le priorità di cura, ma anche identificando chi potrebbe aver bisogno di terapie intensive o di degenze più lunghe, permettendo quindi un'allocazione più efficiente del personale, delle attrezzature e dei posti letto. È il caso **dell'ospedale pubblico Parc Taulí, in Spagna (Catalogna)**.

Un altro esempio di AI applicata alla fase di triage è un **modello progettato per diagnosticare e assistere nella gestione delle embolie polmonari** che, in molti ospedali, ha permesso una **riduzione dei tempi di risposta (TAT)**,

di trattamento e di attesa. Presso il Region Halland Health System in Svezia, ad esempio, il TAT è diminuito da 24 ore a circa 40 minuti, mentre il tempo di trattamento è sceso da 28 ore a quasi un'ora.

Le **potenzialità in ambito diagnostico** sono ben note (in uno studio condotto negli USA, l'AI è riuscita ad individuare col 91% di accuratezza il tumore al collo dell'utero; l'occhio umano col 69%), ma ci sono dei **risultati promettenti anche a livello terapeutico**, soprattutto per la cura del cancro. In tal senso, è rilevante uno **studio retrospettivo, condotto dall'Istituto oncologico olandese**, che ha valutato un algoritmo AI sviluppato per identificare possibili *pattern* all'interno di immagini mediche, che potrebbero fungere da biomarcatori utili a prevedere la risposta dell'organismo al trattamento. Sono state analizzate più di 1000 lesioni tumorali di 203 pazienti con melanoma avanzato e carcinoma polmonare trattati con l'immunoterapia. L'algoritmo ha previsto la risposta fisica alla cura con un'accuratezza complessiva del 76% e un miglioramento del 24% dei tassi di sopravvivenza a 1 anno.

Con l'AI, poi, **strumenti diagnostici e terapeutici avanzati** diventerebbero **accessibili anche alle popolazioni svantaggiate**, rendendo **l'accesso alle cure più equo**. In **Sudafrica**, per esempio, 6 algoritmi sono riusciti a prevedere con elevata accuratezza la recidiva del cancro del colon-retto, aumentando i tassi di sopravvivenza. Oppure in **Etiopia**, è stato possibile ridurre le tempistiche necessarie a identificare il tipo di leucemia da 30 minuti a meno di un minuto, migliorando al contempo l'accuratezza dal 70% al 97%.

Insomma, le promesse sono tante, le potenzialità altrettante ma, secondo il [World Economic Forum](#), nelle corsie la diffusione dell'AI è ancora «al di sotto della media» rispetto ad altri settori. C'è ancora molta strada da fare.

*il [link](#) al report

...a patto di non diventarne dipendenti

Immaginate un medico che si abitui a prendere decisioni con l'AI fino a diventarne dipendente. Oppure un **ospedale paralizzato dalla temporanea indisponibilità dei sistemi d'Intelligenza artificiale**. Non è fantascienza, ma la rappresentazione delle preoccupazioni che disturbano i sonni dei ricercatori, vista anche la rapidità con cui gli LLM stanno conquistando spazi sempre maggiori.

Un editoriale dell'agosto scorso su [The Lancet Digital Health](#) riferiva che Microsoft ha annunciato di lavorare a **un'AI generativa molto più performante dei medici**: di 10 casi complessi tratti dal *New England Journal of Medicine*, l'Intelligenza artificiale ne ha risolti 8, contro i due dei medici in carne e ossa. L'utilizzo clinico è prematuro, ma altre AI generative vengono già impiegate dai sistemi sanitari. **In centinaia di ospedali cinesi dal 2023 si utilizza DeepSeek**, open source e low cost, per la diagnosi, per supportare decisioni e per il management ospedaliero. La **Gran Bretagna** ha invece annunciato che utilizzerà – prima al mondo – un **sistema di alert in grado di analizzare i dati dell'ospedale e identificare i rischi di morte in utero**, di morte neonatale e lesioni cerebrali. Se migliorano efficienza e accesso alle competenze – sottolinea l'articolo –, tutto ciò non sembra privo di rischi sostanziali per la sicurezza del paziente e la qualità delle cure. È necessario **integrare l'educazione clinica e il mantenimento delle competenze** all'interno della governance dell'Intelligenza artificiale, così che i clinici mantengano il timone e ci sia il tempo per costruire valide norme.

È ancora *The Lancet*, questa volta nella sezione [Gastroenterology & Hepatology](#), a portare dati sulla potenziale **perdita di competenze (deskilling)** e la **ridotta capacità di acquisirne di nuove (upskilling inhibition)** per i sanitari abituati a utilizzare l'AI. Uno studio ricorda che, nonostante l'AI migliori il riconoscimento di adenomi nel corso di indagini endoscopiche, il suo utilizzo non si traduce automaticamente in cure migliori. Krzysztof Budzyń e colleghi – dopo aver esaminato i risultati di quattro centri di endoscopia in Polonia su oltre 1.400 pazienti -

sottolineano che i **medici abituati ad avvalersi dell'AI, quando devono farne a meno, riconoscono il 6% di lesioni in meno**, dal 28,4% al 22,4%.

Al tema, trasversale a più specialità cliniche, è dedicata una revisione di 62 papers – disponibile su SSRN pre review, prima autrice Chiara Natali, Università di Milano Bicocca – secondo cui il ricorso all'AI minerebbe alcune fra le principali competenze dei professionisti sanitari: l'esame del paziente, il ragionamento diagnostico, la comunicazione medico-paziente

A rischio *deskilling* sono soprattutto i giovani medici. Tradizionalmente, tirocinanti o neolaureati vengono inseriti in un contesto in cui, affiancando colleghi più esperti, possono mettersi gradualmente alla prova con casi sempre più complessi. «Quando i sistemi di AI forniscono costantemente soluzioni, i tirocinanti rischiano di perdere opportunità fondamentali per sviluppare **acume diagnostico, capacità di problem solving e fiducia nel proprio giudizio indipendente**. A lungo termine, questa inibizione potrebbe portare alla formazione di **una generazione di clinici meno preparati a operare senza il supporto dell'AI**».

Secondo gli autori alcune competenze potrebbero andare **perse in modo irreversibile** e questo può tradursi in un **aumento degli errori medici**. Venendo meno l'esperienza, il responso dell'AI non verrebbe messo in discussione e il ruolo degli umani resterebbe puramente formale.

Non solo: se l'AI si sviluppa a partire da dati consolidati, **il sapere rischia di cristallizzarsi** e di perdere la possibilità di progredire. Si creerebbe anche un ***deskilling* dei sistemi**: addio alle decisioni condivise dal team, alla collaborazione professionale e alla trasmissione non codificata del sapere. In poche parole, alla medicina come la conosciamo...

E intanto tra i medici spunta una domanda: cosa dico al paziente sull'uso dell'AI?

«Cosa dobbiamo dire ai pazienti?». Con il rapido sviluppo delle **applicazioni sociosanitarie dell'AI**, sono sempre di più i medici che si pongono questa domanda, spesso senza trovare una risposta (anche per via della **mancanza di un quadro normativo specifico**).

Ad affrontare questo tema è [uno studio pubblicato su JAMA](#), l'organo ufficiale dell'American Medical Association, che propone un **framework di riferimento** a cui i professionisti possono affidarsi per capire cosa sia più opportuno fare a seconda delle circostanze. **Le opzioni sono tre: notificare** direttamente ai pazienti l'uso dell'AI, **chiedere il loro consenso esplicito** oppure **nessuna delle due** e non dire niente. Come si fa a decidere? In teoria, «quando un'organizzazione sanitaria decide di usare uno strumento AI, deve stabilire se ha l'obbligo etico di informare i pazienti o ottenere il loro consenso», [spiega Jamie M. Mello](#), bioeticista della Stanford Law School che ha condotto lo studio. In pratica, però, non è così.

Il modello proposto nel report suggerisce una **strategia basata su due considerazioni: qual è la probabilità che l'AI possa causare danni ai pazienti e che possibilità hanno questi ultimi di esercitare, per davvero, il loro potere decisionale** nel contesto dell'assistenza sanitaria?

Se il **rischio di danni e l'opportunità di agire sono elevati, l'obbligo di informativa è massimo**: ove possibile si chiede il consenso, altrimenti si avvisano i pazienti. Nel caso in cui, invece, entrambi questi parametri siano assenti – o per lo meno molto bassi –, lo studio suggerisce la terza opzione: non fare nulla.

Il paper propone anche alcuni **casi concreti**. Per esempio, **se ad eseguire un'operazione chirurgica fosse un robot** (non autonomo e guidato dall'AI), bisognerebbe chiedere il **consenso** poiché c'è una reale possibilità che vengano commessi errori e il paziente potrebbe preferire un'altra tipologia di intervento. Oppure, **quando si usa l'AI per rispondere alle mail dei pazienti, questi devono essere informati**, in modo che possano chiedere

spiegazioni su risposte strane o imprecise. Invece, **non si deve fare nulla se l'AI viene usata per riassumere i risultati di un esame diagnostico**: la probabilità di commettere un errore e fare del male all'assistito è molto bassa, perché l'AI si basa sulle indicazioni fornite dal radiologo stesso.

C'è ancora **molta gente che non si fida dell'AI** e quindi è contraria al suo utilizzo. Secondo uno studio condotto negli Stati Uniti, **l'80% circa degli adulti non crede che possa migliorare l'assistenza medica** e solo un terzo si fida del suo uso responsabile nei sistemi sanitari. In ogni caso, il **63% vuole essere avisato** del suo impiego. Informare i pazienti è importante ma, affinché l'uso dell'AI venga accettato, è altrettanto rilevante «ottenere la loro fiducia». Per costruirla, lo studio suggerisce di **comunicare in modo chiaro, semplice e sintetico** le modalità con cui l'organizzazione usa l'AI per «fornire un'assistenza migliore, più sicura e più efficiente».

Per approfondire

*il [link](#) al paper

[R. Corcella, AI in corsia: il paziente deve sapere quando viene usata? E come? Cosa chiedere al medico | Corriere della Sera, 31 agosto 2025](#)

[Maintaining Safety and Trust When Patients Engage Google: A Conversation With Dr Michael Howell | JAMA](#)

Focus



Il Tribunale di Torino riconosce la lite temeraria per un ricorso redatto con sistemi di IA *di Marco Bassini*

Con sentenza del 16 settembre 2025, il Tribunale di Torino ha affrontato un caso di particolare rilievo per **l'impatto dei sistemi di AI generativa nella redazione degli atti giudiziari**. Il giudice ha **rigettato un ricorso ritenuto infondato perché redatto con il supporto dell'AI**, condannando la parte ricorrente anche per lite temeraria ai sensi dell'articolo 96, comma 3, del Codice di procedura civile.

Il ricorso contestava una serie di avvisi di addebito e un'ingiunzione di pagamento per contributi previdenziali, ma **le eccezioni sollevate** sono risultate **prive di fondamento**, nonché generiche e **senza alcun riferimento concreto ai singoli atti impugnati**. La giudice ha sottolineato come le argomentazioni proposte si limitavano a «un coacervo di citazioni normative e giurisprudenziali astratte, prive di ordine logico e in larga parte inconferenti, senza allegazioni riferibili in concreto alla situazione oggetto del giudizio». Tale modalità di redazione ha indotto il tribunale a ritenere che il **ricorso fosse stato predisposto in modo negligente**, se non addirittura in mala fede.

Elemento centrale della decisione è stato **l'impiego dell'AI** per la scrittura del ricorso, **senza una reale attività di verifica**, interpretazione e contestualizzazione da parte del difensore. Secondo il Tribunale, **l'uso dell'AI non può mai sostituire il dovere di diligenza richiesto al difensore nella predisposizione degli atti**, soprattutto laddove questi contengano argomentazioni standardizzate e scollegate dalla fattispecie concreta. L'adozione acritica di contenuti generati da strumenti automatizzati non solo mina la qualità del processo, ma espone il professionista e la parte a gravi conseguenze, come appunto la condanna per responsabilità processuale aggravata.

Il giudice ha quindi condannato la ricorrente al pagamento delle spese legali, a una somma di 500 euro in favore di ciascuna delle controparti e a ulteriori 500 euro da versare alla Cassa delle Ammende. La **sentenza** rappresenta un precedente significativo, poiché **sottolinea la necessità di un uso consapevole e responsabile delle tecnologie emergenti nel processo civile**, richiamando con forza il ruolo centrale dell'attività critica, personale e professionale dell'avvocato.

Non si tratta, tuttavia, di un caso isolato. Pochi mesi prima, il Tribunale delle Imprese di Firenze si era pronunciato su una **vicenda analoga: un avvocato aveva inserito nella propria memoria difensiva sentenze completamente inventate da ChatGPT**, salvo poi ricondurne la responsabilità a una collaboratrice che aveva usato il sistema per ricerche giurisprudenziali. **Il tribunale fiorentino** (con ordinanza del 14 marzo 2025), pur censurando l'omesso controllo da parte della difesa, **ha escluso l'applicazione della sanzione ex art. 96 c.p.c.**, riconoscendo **l'assenza di mala fede** e rilevando che i **riferimenti falsi non avevano inciso sulla sostanza della linea difensiva**.

Il confronto tra i due casi evidenzia come, a fronte dell'uso dell'AI nei procedimenti giudiziari, **ciò che può fare la differenza** non sia tanto lo strumento in sé, quanto **il grado di consapevolezza, verifica e responsabilità professionale** che vengono esercitati.

Normativa

Legge AI e responsabilità della PA: la compliance 231/2001 per gli enti pubblici economici
a cura di *Annalisa Negrelli*

La **Legge 132/2025** recepisce l'AI Act europeo, introducendo **nuove regole sull'uso dell'AI nella Pubblica amministrazione**. L'art.14 pone l'accento su una **serie di obiettivi**: incrementare l'efficienza, ridurre i tempi dei procedimenti, aumentare la qualità e la quantità dei servizi erogati – con un focus sulla trasparenza nell'uso di questi strumenti e sull'imprescindibile potere decisionale dell'uomo. Si sottolinea anche **l'esigenza di adottare «misure tecniche, organizzative e formative»** per garantire un uso responsabile dell'AI. Una trasformazione, dunque, che parte dalle persone in quanto impatta sui processi interni alle amministrazioni, su ruoli e competenze.

In armonia con la normativa europea, la legge 132/2025 introduce **nuove fattispecie penali e aggravanti legate all'uso dell'AI**, ampliando i reati presupposto. Tra le varie novità, applicabili in particolare agli enti pubblici che esercitano attività economiche (enti pubblici economici, società a partecipazione pubblica, società in house), si evidenzia **l'obbligatorietà dell'adeguamento dei Modelli di organizzazione e gestione previsti dalla legge 231/2001 a queste nuove disposizioni sanzionatorie**, a partire da una revisione delle regole, delle responsabilità e dei controlli sulla gestione dei sistemi intelligenti così da prevenire rischi e garantire un uso etico delle tecnologie. Ne derivano **impatti significativi sui modelli organizzativi**, che dovranno integrare **misure di controllo sui processi automatizzati**.

Emergono **concetti innovativi**, come la **“colpa artificiale”**, che ridefiniscono la governance – richiedendo un approccio interdisciplinare tra diritto, tecnologia e compliance per gestire eventuali rischi – e gli assetti di responsabilità, in uno scenario in cui si intersecano indubbi vantaggi competitivi, ma anche nuovi rischi legati alla possibilità che tali strumenti vengano usati, magari in modo inconsapevole, per la commissione di reati.

I reati legati all'uso dell'AI e le nuove aggravanti

La nuova legge italiana sulla AI è intervenuta **in modo incisivo sul fronte sanzionatorio** prevedendo **l'introduzione di un'aggravante comune** applicabile a tutti i reati e, di conseguenza, anche a quelli presupposto della responsabilità amministrativa dell'Ente.

L'art. 26 (comma 1, lett. a) ha aggiunto all'art. 61 il nuovo comma 11 *decies*, con cui si introduce «una nuova aggravante quando il reato è commesso mediante l'impiego di sistemi di AI allorché, per la loro natura o modalità d'uso, questi abbiano costituito un mezzo insidioso, abbiano ostacolato le attività di difesa (pubblica o privata), oppure abbiano aggravato le conseguenze del reato».

In tema di circostanze del reato, **viene prevista un'aggravante specifica, qualora il fatto venga commesso mediante sistemi di AI, per alcune fattispecie di reato presupposto** della responsabilità dell'Ente, quali **l'aggiotaggio** (art. 2637 c.c.) e la **manipolazione del mercato** (art. 185 TUF).

La Legge introduce il **nuovo reato**, di cui all'art. 612 *quater* c.p., di **«illecita diffusione di contenuti generati o alterati con sistemi di Intelligenza artificiale»**. La norma punisce con la reclusione da uno a cinque anni «chiunque cagioni un danno ingiusto ad una persona, cedendo, pubblicando o altrimenti diffondendo, senza il suo consenso, immagini, video o voci falsificati o alterati mediante l'impiego di sistemi di Intelligenza artificiale e idonei a indurre in inganno sulla loro genuinità», i cosiddetti *deepfake*.

Aggiornamento del modello 231 e formazione sull'uso consapevole dell'AI

Gli enti pubblici economici devono quindi **aggiornare il Modello 231 e le procedure di gestione del rischio** al fine di: **intercettare i nuovi rischi di illecito** connessi all'impiego dell'AI; **integrare misure specifiche di controllo** sull'uso degli algoritmi e sull'accesso ai sistemi (anche ai fini dell'impiego di dati di persone fisiche il cui trattamento è sottoposto alle tutele di cui al **Regolamento UE 2016/679 – GDPR**- e di cui al Codice privacy); **implementare la formazione del personale** sull'uso sicuro e conforme dell'AI.

Su quest'ultimo aspetto, è intervenuto **l'AI ACT** con una **disposizione specifica** (di cui all'art. 4 del Regolamento Ue), rubricata **«alfabetizzazione in materia di AI»**. La norma prevede che le società che forniscono e sviluppano sistemi di AI, ma anche gli enti che li usano, debbano adottare «misure per garantire un livello sufficiente di alfabetizzazione in materia di AI del loro personale». Questa norma ha un impatto anche sulla compliance 231, al fine di assicurare un uso virtuoso e consapevole dell'AI nei contesti pubblici e privati, attraverso un'adeguata formazione ai dipendenti.

Per approfondire

[F. Sbisà e C. Di Lorenzo, L'impatto dell'intelligenza artificiale sulla disciplina del d.lgs. 231/2001 | Altalex ilQG, 8 agosto 2025](#)

Applicazioni alla Pubblica Amministrazione

ITALIA

M-Apperò- Brescia

Sistema sensoristico di monitoraggio di urban mobility che, attraverso la sollecitazione delle sospensioni delle auto, raccoglie dati significativi per i mezzi e i territori su cui transitano, favorendo così l'innovazione nei contesti urbani e sostenendo la manutenzione predittiva.

[M-Apperò – lobo](#)

Progetto AI Lis e LIS2Speech

Sviluppato dall'Università di Siena, si tratta di un avatar in grado di parlare la lingua dei segni (Lis) nei portali della PA, delle banche e delle strutture sanitarie. La particolarità di questo assistente virtuale è un software che utilizza una webcam capace di percepire l'utente, analizzare ciò che esprime attraverso la lingua dei segni e fornire le risposte adeguate riproducendole sempre in Lis.

[A. Piemontese, Tradurre la lingua dei segni con l'AI, l'idea di una startup sociale | Wired Italia, 22 giugno 2024](#)

[L'AI al servizio delle tecnologie assistive: il progetto LIS2Speech per la traduzione della lingua dei segni | Orbyta](#)

UNIONE EUROPEA

AI Polonia

Una mappatura di più di 100 buone pratiche e di applicazioni di AI alla pubblica amministrazione in Polonia, dalla modellazione intelligente nello sviluppo delle smart city, all'illuminazione, dal monitoraggio delle acque ai sistemi di videosorveglianza.

[Mappa dell'Innovazione- Ministero degli Affari Digitali- Portale Gov.pl](#)

MONDO

Open HMRC- UK

Nel Regno Unito un assistente virtuale supporta il rinnovo, il controllo e la verifica dei crediti d'imposta.

[HMRC online services](#)

AI in pillole

“Black Box AI” vs. “White Box AI”. Che cosa si intende e come funzionano?

a cura di Annalisa Negrelli

La “Black Box AI” è un sistema AI il cui funzionamento interno non è visibile, o rimane sconosciuto ai suoi utenti e talvolta persino ai suoi sviluppatori. Gli utilizzatori possono vedere gli input e gli output del sistema, ma non sono in grado di conoscere i meccanismi interni con cui lo strumento di AI ha generato tali output.

Questi modelli possono originarsi in due modi: gli sviluppatori trasformano i modelli AI in scatole nere di proposito, oppure diventano scatole nere in modo naturale, come sottoprodotto del loro addestramento. Nel primo caso, viene oscurato il funzionamento interno prima lancio sul mercato, con l'intento di proteggere la proprietà intellettuale e mantenere segreti sia il codice sorgente, che il processo decisionale. Nel secondo caso, che si verifica sempre più spesso con tecnologie AI molto avanzate, i creatori non oscurano i meccanismi in modo intenzionale, ma i sistemi di deep learning che alimentano i modelli sono così complessi che persino gli stessi programmatori non sono in grado di spiegare esattamente che cosa succeda al loro interno. Sono le cosiddette “scatole nere organiche”.

Molti dei modelli di apprendimento automatico oggi disponibili, come ChatGPT di OpenAI e Lama di Meta, sono *black box AI*.

Ma perché è così difficile conoscere ciò che avviene al loro interno? Questi modelli di AI generativa vengono addestrati su enormi set di dati attraverso processi di deep learning e si basano su reti neurali complesse di cui si servono per rispondere a comandi in linguaggio naturale, per risolvere nuovi problemi e per creare contenuti originali. Il problema della scatola nera, quindi, non si può risolvere semplicemente usando strumenti di AI tradizionali e più spiegabili perché non sarebbero così potenti e così flessibili come invece è l'AI generativa.

Sebbene la *black box AI* possa fornire risultati impressionanti, gli utenti non sempre riescono a fidarsi degli output per via della mancanza di trasparenza: non è facile convalidare i risultati di un modello se non si conoscono i meccanismi che li hanno generati. Inoltre, dietro all'opacità di un modello a scatola nera, si possono nascondere una serie di problemi come la vulnerabilità della sicurezza informatica, i pregiudizi e le violazioni della privacy.

Da questa prospettiva, la *black box AI* pone alcune sfide:

- **Riduzione dell'attendibilità negli output del modello:** non solo gli utenti non sanno come un modello a scatola nera generi gli output, ma può anche capitare che i modelli *black box* giungano alle conclusioni giuste per il motivo sbagliato (all'insaputa dei loro utenti). Questo fenomeno è chiamato "effetto Clever Hans", dal nome dell'esperimento condotto su un cavallo che apparentemente era in grado di contare e fare semplici calcoli aritmetici battendo lo zoccolo, ma che in realtà si basava su sottili segnali dal linguaggio del corpo del suo padrone per capire quando fosse il momento di smettere di calpestare. L'effetto Clever Hans può avere gravi conseguenze quando i modelli vengono applicati a campi sensibili, come ad esempio quello dell'assistenza sanitaria.
- **Difficoltà nella regolazione delle operazioni del modello:** se un modello *black box* prende decisioni sbagliate, o produce costantemente output imprecisi o dannosi, può diventare difficile modificare il modello per correggerne il comportamento, perché non possono essere individuati gli errori in cui incorre. Questo problema rappresenta una sfida significativa in particolare nel campo dei veicoli autonomi, in cui gli sviluppatori addestrano sofisticati sistemi di AI per prendere decisioni di guida in tempo reale che, se errate, possono risultare fatali.
- **Problemi di sicurezza:** i sistemi di AI generativa *black box* sono vulnerabili agli attacchi al software e al cosiddetto avvelenamento dei dati, che possono modificare il comportamento del modello a sua insaputa.
- **Preoccupazioni etiche:** ogni strumento di AI può riprodurre i pregiudizi umani se presenti nei dati di addestramento o nel design. Con i modelli a scatola nera può essere molto difficile individuarne l'esistenza o le cause.
- **Non conformità normativa:** alcune normative, come l'EU AI Act e il California Consumer Privacy Act, stabiliscono regole sull'uso dei dati personali sensibili con strumenti decisionali basati sull'AI. Con la

black box AI può essere difficile per un'organizzazione sapere se il modello è conforme o dimostrarlo in caso di audit.

In risposta a queste sfide, aziende ed enti di ricerca (come **Anthropic** e **l'Università dell'Illinois**) stanno lavorando per rendere gli algoritmi di AI più "spiegabili", cercando di **svelare i misteri della *black box*** usando le tecniche di **Explainable AI** (vedi il [numero precedente](#) di *PoliS-AI News*).

L'AI "**white Box**" o "**a scatola bianca**", chiamata anche AI spiegabile (XAI) o AI a scatola di vetro, è l'opposto di quella a scatola nera. Si tratta di un sistema di Intelligenza artificiale con **funzionamento interno trasparente**: gli utenti capiscono come l'AI acquisisca i dati, li elabori e arrivi ad una conclusione.

I modelli *white box* **semplificano l'attendibilità e la convalida dei risultati**, oltre a modificare i modelli per correggere gli errori e regolare le prestazioni. Ma non è facile trasformare ogni AI in una scatola bianca. I modelli tradizionali, spesso, possono essere resi trasparenti condividendo il loro codice sorgente. Ma i modelli di machine learning più sofisticati sviluppano i propri parametri attraverso algoritmi di deep learning e il semplice fatto di **avere accesso alle loro architetture non sempre è sufficiente per capire cosa stiano facendo**.

Altri ricercatori hanno sviluppato tecniche per spiegare come i modelli arrivino a conclusioni specifiche. Ad esempio, la **spiegazione indipendente dal modello interpretabile locale (LIME)** è un processo che utilizza un modello di apprendimento automatico separato per analizzare le relazioni tra gli input e gli output di una scatola nera, con l'obiettivo di identificare le caratteristiche che potrebbero influenzare gli output del modello.

In sintesi, un riepilogo delle differenze tra l'AI a scatola nera e l'AI a scatola bianca:

- L'AI a scatola nera è spesso più precisa ed efficiente dell'AI a scatola bianca.
- L'AI a scatola bianca è più facile da capire dell'AI a scatola nera.
- I modelli a scatola nera includono modelli di foresta casuale e di potenziamento che sono di natura altamente non lineare e più difficili da spiegare.
- Nell'AI white box è più facile da eseguire il debug e la risoluzione dei problemi grazie alla sua natura trasparente e interpretabile.
- Alcuni tecnicismi per addetti ai lavori: l'albero lineare, l'albero decisionale e l'albero di regressione sono tutti modelli di intelligenza artificiale white box.

Per approfondire:

[M. Kosinki, *What is black box artificial intelligence \(AI\)?* | IBM](#)

[L. Mischtelli, *Interpretabilità dell'IA: ecco i primi passi per risolvere un grosso problema* | Agenda Digitale, 14 giugno 2024](#)

[Navigare nel dibattito sull'intelligenza artificiale nella scatola nera nel settore sanitario | TechTarget, 1 maggio 2024](#)

[K. Kelley e B. St. George, *Risolvere il problema della scatola nera dell'IA attraverso la trasparenza* | TechTarget, 16 agosto 2021](#)

Notizie

[P. L. Pisa, *Deloitte rimborsa il governo australiano dopo gli errori di un report scritto con l'IA* | La Repubblica, 7 ottobre 2025](#)

[G. Rusconi, *L'ONU entra in gioco nella governance dell'AI. Ecco cosa sta facendo* | Il Sole 24 Ore, 7 ottobre 2025](#)

[C. La Via, *Intelligenza artificiale al volante delle flotte* | Il Sole 24 Ore, 7 ottobre 2025](#)

[V. Alvich, *Intelligenza artificiale in classe, istruzioni per l'uso: «La rivoluzione parte proprio dai dirigenti scolastici»* | Corriere della Sera, 6 ottobre 2025](#)

[M. Carmignani, *California, perché la nuova legge sull'AI apre allo scontro con le Big Tech* | Agenda Digitale, 3 ottobre 2025](#)

[F. Santelli, *Bezos: "L'IA è una bolla ma darà benefici". Elkan: "Nel tech c'è anche l'Europa"* | La Repubblica, 3 ottobre 2025](#)

[E. Spagnuolo, *«Diamo dei limiti all'AI». L'appello di oltre 200 scienziati, politici e premi Nobel: «Rischiamo di perdere il controllo»* | Corriere della Sera, 2 ottobre 2025](#)

[S. Tirrito, *Stellantis e Mistral, nuova partnership su auto e Ai: l'annuncio alla Italian Tech Week* | La Stampa, 1 ottobre 2025](#)

[C. Crescenzi, *Grokopedia, l'enciclopedia competitor di Wikipedia annunciata da Elon Musk* | Wired, 1 ottobre 2025](#)

[G. Finocchiaro, *La legge italiana va verso la semplificazione, soprattutto in ambito sanitario* | Il Sole 24 Ore, 1 ottobre 2025](#)

[Z. Schiffer, L. Matsakis, *Sora 2, OpenAI è pronta a lanciare un'app social che ricorda da vicino TikTok* | Wired, 30 settembre 2025](#)

[D. Manca, *Lavoro, tasse e intelligenza artificiale: il dinamismo che manca all'Italia* | Corriere della Sera, 29 settembre 2025](#)

Commenti

[C. De Gregorio, *Certi romanzi tutti uguali e un pensiero molesto sull'intelligenza artificiale* | La Repubblica, 7 ottobre 2025](#)

[A. Butti, *Intelligenza artificiale in sanità: ecco come coniugare innovazione, empatia e inclusione* | Il Sole 24 Ore, 6 ottobre 2025](#)

[E. Mazza, *Musica e IA, perché è importante che al centro sia sempre la creatività* | La Repubblica, 5 ottobre 2025](#)

[C. Dell'Acqua, *Come educare al pensiero nell'era dell'Intelligenza artificiale* | Corriere della Sera, 4 ottobre 2025](#)

[The Ezra Klein Show, *Are You Playing the Technology or Is the Technology Playing You?* | New York Times, 3 ottobre 2025](#)

[U. Bertelè, *AI, ma è bolla finanziaria? Facciamo chiarezza* | Agenda Digitale, 3 ottobre 2025](#)

[G. C. Wong, 'My son genuinely believed it was real': Parents are letting little kids play with AI. Are they wrong? | The Guardian, 2 ottobre 2025](#)

[A. Puliafito, L'intelligenza artificiale può semplificare l'Inps? | Internazionale, 1 ottobre 2025](#)

[P. Benanti e S. Maffettone, Se ci affezioniamo all'intelligenza artificiale | Corriere della Sera, 1 ottobre 2025](#)

[G. R. Anthis, It's time to prepare for AI personhood | The Guardian, 30 settembre 2025](#)

[P. Raimondi, UN: la governance dell'AI e armi autonome letali | Rivista AI, 30 settembre 2025](#)

[W. Veltroni, Tilly Norwood, l'attrice che non esiste perché creata con l'intelligenza artificiale. Ma è già stata ingaggiata da un'agenzia di talenti | Corriere della Sera, 30 settembre 2025](#)

[G. A. Fowler, I discovered ChatGPT's best new feature: Quitting things for you | The Washington Post, 29 settembre 2025](#)

Corsi, convegni e pubblicazioni

Corsi

[Luiss Business School, Le tecnologie digitali e di AI per la transizione circolare | dal 13 ottobre 2025](#)

[24 Ore Business School, Master AI e legal compliance | 17 ottobre 2025](#)

[RCS Academy Business School, AI for Business | 22 ottobre 2025- 18 febbraio 2026](#)

Eventi e convegni

[Giunti Psicologia, Psicologia.io & Perlab, Empatia e Intelligenza Artificiale | 7-8 novembre 2025](#)

[Fondazione CRUI, Polimi Graduate School of Management e METID, L'uso dell'AI generativa nella didattica universitaria | 10-18-27 novembre 2025](#)

[Polimi School of Management, Intelligenza Artificiale nelle amministrazioni pubbliche | 19 novembre 2025](#)

Pubblicazioni

[L. Franzese e F. Recanati, Conversazioni su etica e intelligenza artificiale | Giappichelli, 2025](#)

[E. Yudkowsky e N. Soares, If Anyone Builds It, Everyone Dies | Bodley Head, 2025](#)

[E. Bocciolesi e A. De Lucia, Umanità, intelligenza artificiale e sostenibilità | Edizioni Scientifiche Italiane, 2025](#)

Link attivi al 10 ottobre 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)