



MIM

Ministero dell'Istruzione
e del Merito



SICURNAUTI

LA TUA GUIDA NELLA GALASSIA DIGITALE

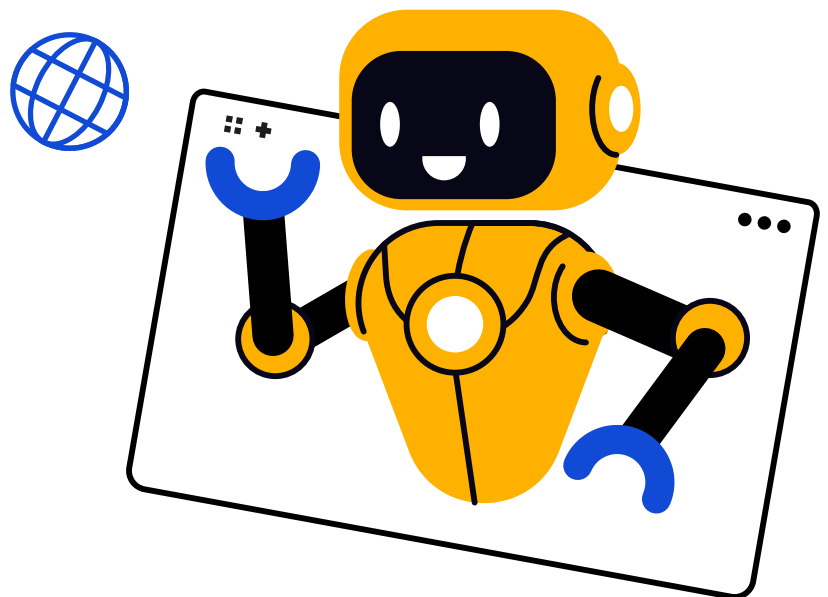


L'INTELLIGENZA ARTIFICIALE IN SICUREZZA

www.unica.istruzione.gov.it

INDICE DEI CONTENUTI

1	INTRODUZIONE: COS'È L'INTELLIGENZA ARTIFICIALE	3
2	CARATTERISTICHE PRINCIPALI	3
3	RISCHI LEGATI AD UN USO IMPROPRIO DELL'IA	5
4	CASI REALI DI INCIDENTI DOVUTI AD UN USO MALEVOLO DELL'IA	6



INTRODUZIONE: COS'È L'INTELLIGENZA ARTIFICIALE

L'**Intelligenza Artificiale (IA)**, frutto dell'integrazione di discipline come informatica, matematica, statistica e neuroscienze, rappresenta una delle **innovazioni più significative e affascinanti del nostro tempo**. Negli ultimi dieci anni, il suo sviluppo ha conosciuto un'**accelerazione straordinaria** grazie ai progressi del *deep learning*, delle reti neurali sempre più sofisticate e alla disponibilità di grandi quantità di dati. Questi fattori hanno reso l'IA una **presenza costante nella vita quotidiana e professionale**, capace di trasformare in profondità i nostri comportamenti, i processi produttivi e i modelli decisionali.

Secondo l'**Osservatorio Artificial Intelligence del Politecnico di Milano**, il mercato dell'IA ha registrato una **crescita esponenziale** a partire dal 2018, con un incremento del **58% tra il 2023 e il 2024**, fino a raggiungere un valore di **1,2 miliardi di euro solo in Italia**.

Oggi l'Intelligenza Artificiale, infatti, è ormai **ampiamente diffusa** e, grazie alla combinazione di **dati, algoritmi complessi e potenza di calcolo**, si è affermata come una tecnologia estremamente efficace, sempre più integrata nella società e in grado di **supportare le persone, migliorare i servizi e semplificare attività complesse**.

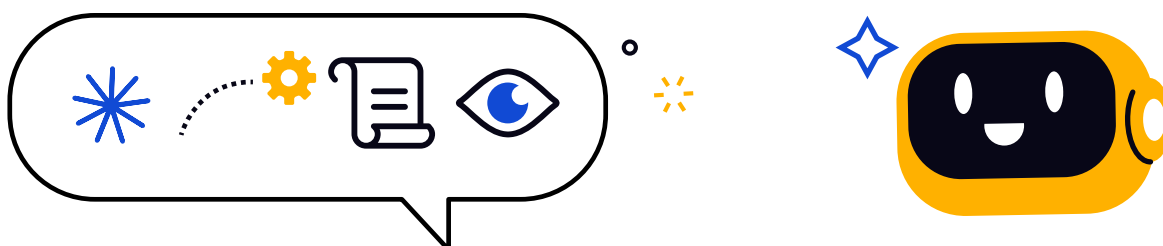
L'evoluzione dell'Intelligenza Artificiale non rappresenta soltanto una conquista tecnologica, ma un vero e proprio **linguaggio del progresso umano**, capace di unire competenze, valori e visioni per costruire un mondo più aperto, sostenibile e intelligente. Tuttavia, di fronte a questa straordinaria opportunità, è essenziale adottare un approccio basato su **attenzione, responsabilità e consapevolezza**.

Perché i benefici dell'IA siano reali e duraturi, occorre perciò promuoverne un utilizzo fondato sui **principi etici**, sulla **trasparenza** e sulla **protezione dei dati personali**, nel rispetto delle fonti normative e regolamentari in materia, nonché degli orientamenti e delle buone pratiche riconosciute a livello nazionale ed europeo.

In questa direzione si è mossa l'Unione europea, che il 13 giugno 2024 ha approvato il **Regolamento (UE) 2024/1689 (AI Act)**, il quale introduce un quadro normativo comune in materia di progettazione, implementazione e utilizzo dei sistemi di Intelligenza Artificiale, con l'obiettivo di assicurare un livello uniforme di protezione dei diritti e delle libertà degli interessati, nonché di promuovere la diffusione di sistemi di Intelligenza Artificiale antropocentrici e affidabili, concepiti come strumenti a supporto degli individui.



A livello nazionale, la Legge 23 settembre 2025, n. 132, (Legge sull'IA), in armonia con l'AI Act, ha introdotto norme volte a garantire un impiego etico, trasparente e responsabile dei sistemi di IA. Nel rispetto delle fonti suindicate, al fine di incentivare un uso corretto e antropocentrico dell'IA in ambito scolastico e supportare un'applicazione efficace, sicura e inclusiva della stessa nei processi educativi, il Ministero dell'Istruzione e del Merito ha adottato, con Decreto Ministeriale 9 agosto 2025, n. 166, le «**Linee guida per l'introduzione dell'Intelligenza Artificiale nelle Istituzioni scolastiche**», le quali forniscono un quadro di riferimento strutturato per l'adozione sicura e consapevole delle soluzioni e dei sistemi di Intelligenza Artificiale nelle Istituzioni scolastiche, rivolgendosi ai principali attori del settore, tra cui dirigenti scolastici, personale amministrativo, docenti e studenti.



2 CARATTERISTICHE PRINCIPALI

Quando parliamo di Intelligenza Artificiale facciamo riferimento ad uno strumento in continua evoluzione. Tra le sue caratteristiche fondamentali troviamo:

Capacità di apprendimento: attraverso tecniche di *machine learning*, i sistemi IA analizzano una grande quantità di dati, apprendendo da *pattern* ed esperienze che possono migliorare le proprie prestazioni nel tempo (addestramento).

Elaborazione del linguaggio naturale: le macchine sono in grado di comprendere, interpretare e generare linguaggio umano in modo coerente e comprensibile che migliora l'esperienza dell'utente.

Adattabilità: i sistemi possono modificare il loro comportamento in risposta a nuovi dati o situazioni, migliorando continuamente le prestazioni.

Velocità e profondità di analisi: gli algoritmi IA sono in grado di elaborare dati massivi, individuando correlazioni e offrendo analisi che sarebbero difficilmente raggiungibili dall'uomo.

Autonomia operativa: l'IA può svolgere determinate attività, prendere decisioni e gestire flussi di lavoro in modo automatico e indipendente, in certi casi anche senza supervisione umana.

In definitiva, l'Intelligenza Artificiale non rappresenta semplicemente l'introduzione di una nuova tecnologia, ma il **motore di una trasformazione strategica** che ridefinisce modelli operativi, processi decisionali ed esperienze utente.

La vera opportunità risiede nella capacità di tradurre i suoi benefici in un vantaggio concreto e sostenibile, capace di generare **valore per la società nel suo insieme**.

3 RISCHI LEGATI AD UN USO IMPROPRIO DELL'IA

L'adozione delle soluzioni di Intelligenza Artificiale, nonostante offra notevoli opportunità, può comportare anche **potenziali rischi**. Infatti, con l'evoluzione dell'IA, anche la complessità delle minacce informatiche è aumentata e i cybercriminali ricorrono sempre più spesso a **schemi di phishing potenziati** dall'IA e a **deepfake**, ovvero foto, video e audio creati grazie a software di Intelligenza Artificiale che, partendo da contenuti reali come immagini e audio, riescono a creare, in modo semplice e senza l'utilizzo di particolari risorse, contenuti audio e video distorti incredibilmente simili agli originali.

Un **uso irresponsabile** degli strumenti di Intelligenza Artificiale può avere conseguenze rilevanti, sia a livello individuale che sociale. Tra i principali pericoli troviamo:

Dipendenza da IA e perdita di capacità critica

Con l'uso crescente di tecnologie basate sull'IA, gli utenti potrebbero sviluppare una **dipendenza** dalla **tecnologia** per risolvere **problemi quotidiani**, prendere **decisioni** o rispondere a domande. Questo può **ridurre** la **capacità critica** e il **pensiero indipendente**, in quanto gli utenti potrebbero affidarsi eccessivamente alle soluzioni automatiche proposte dagli algoritmi, senza riflettere o valutare attivamente le opzioni.



Mantieni un ruolo attivo e critico nel processo decisionale, utilizzando l'Intelligenza Artificiale come strumento di supporto e non a sostituzione del pensiero umano.
Verifica sempre le **risposte** fornite dai sistemi automatici, confrontandole con altre fonti.

Violazione della privacy e uso improprio dei dati

L'Intelligenza Artificiale elabora enormi quantità di dati, tra i quali potrebbero rientrare anche **dati personali comuni**, di cui all'articolo 4, paragrafo 1, numero 1, del GDPR, nonché **dati appartenenti** alle categorie di cui agli **articoli 9 del GDPR** (ad esempio, dati sulla salute o opinioni politiche) e **10** (ad esempio, dati relativi a condanne penali e reati) **del GDPR**.

Se questi dati venissero compromessi a causa di vulnerabilità dei sistemi o di attacchi informatici, potrebbero essere esposti a furti, violazioni o utilizzi non autorizzati, mettendo a **rischio la privacy** degli individui.



Ricorda di **non inserire dati personali o riservati**, propri o altrui, in strumenti IA. Quando si interagisce con assistenti virtuali basati su *chatbot* o modelli linguistici avanzati, le informazioni inserite vengono elaborate da sistemi che applicano misure di sicurezza informatica e politiche di gestione dei dati; tuttavia, queste non garantiscono una riservatezza assoluta né un controllo completo sull'utilizzo dei dati. Ecco perché è raccomandato, ove possibile, l'impiego di sistemi di IA all'interno di ambienti controllati, che assicurino il trattamento dei dati esclusivamente in ambito scolastico e sotto il controllo diretto del gestore del sistema.

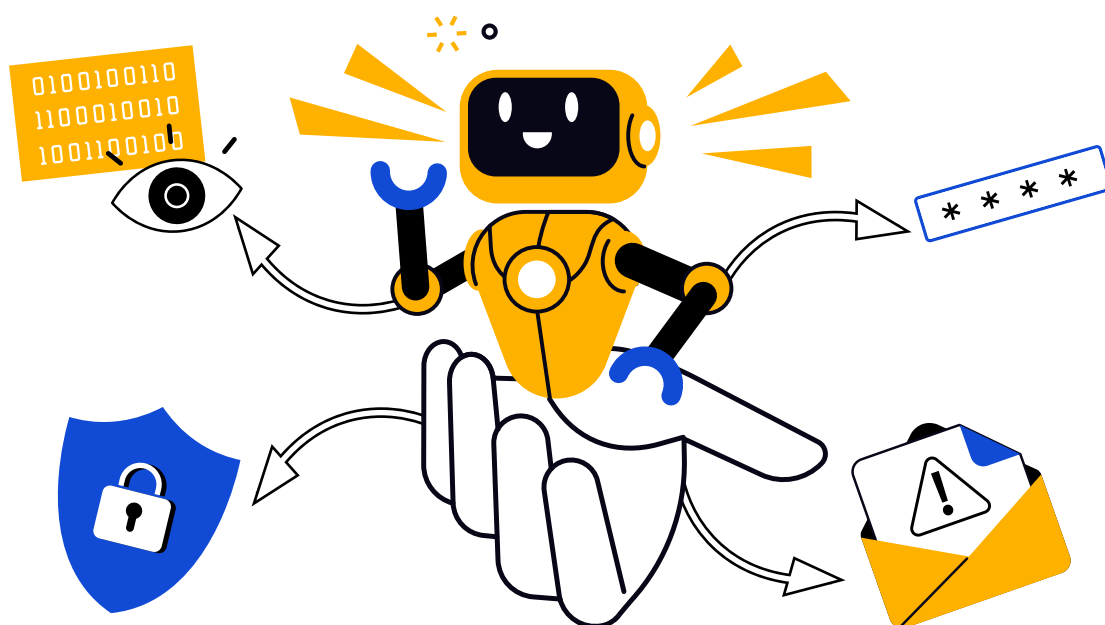
Deepfake e disinformazione

Alcune applicazioni dell'Intelligenza Artificiale, in particolare strumenti per la generazione automatica di contenuti digitali, possono essere utilizzate per creare **news ingannevoli** che, oltre a favorire la **disinformazione online**, possono essere utilizzate attraverso vari canali, inclusi telefonate (con ID chiamante falsificati), e-mail, messaggi di testo e piattaforme di *social media*

per richieste di informazioni sensibili o transazioni finanziarie.

In assenza di regole chiare, controlli efficaci e regolamenti condivisi, questi utilizzi illeciti o non etici rischiano di **compromettere la sicurezza**, minare i processi democratici e indebolire la fiducia nella tecnologia digitale, con conseguenze concrete per la società e per i suoi valori fondamentali.

- ✓ **Verifica sempre l'affidabilità e la provenienza dei contenuti digitali**, in particolare immagini, video ed e-mail, considerando che i messaggi generati con Intelligenza Artificiale possono risultare molto realistici e privi degli errori tipici del *phishing* tradizionale. Inoltre, l'uso di strumenti specializzati per il rilevamento di manipolazioni e contenuti contraffatti aiuta a ridurre il rischio di truffe, disinformazione e attacchi mirati, rafforzando la sicurezza digitale e la consapevolezza nell'utilizzo delle tecnologie.



E-mail di *phishing*

Le e-mail di *phishing* generate con sistemi di Intelligenza Artificiale sono messaggi costruiti per **apparire completamente autentici** e indurre il destinatario a **fornire dati sensibili**, autorizzare operazioni o compiere determinate azioni. L'IA viene sfruttata dai cybercriminali per rendere le campagne ancora più **sofisticate, difficili da individuare e da contrastare**. Grazie ad essa riescono a creare comunicazioni altamente personalizzate, che mimano in maniera quasi completa quelle reali, occultando i consueti indicatori.

- ✓ Per contrastare le e-mail di *phishing* generate con Intelligenza Artificiale è essenziale **verificare con attenzione l'autenticità del mittente** e mantenere un **atteggiamento prudente** verso ogni comunicazione sospetta. È importante trattare con cautela *link*, allegati, URL abbreviati e codici QR, poiché possono celare contenuti malevoli difficilmente individuabili. Inoltre, quando vengono richiesti dati personali o credenziali, è consigliato effettuare la verifica accedendo direttamente dai canali ufficiali, senza utilizzare i collegamenti inclusi nell'e-mail. Adottare queste pratiche e mantenere alta la consapevolezza digitale aiuta a ridurre significativamente il rischio di cadere vittima di campagne di *phishing* evolute.

Vishing (Voice phishing)

I moderni sistemi di Intelligenza Artificiale, partendo da brevi campioni audio, sono in grado di **ricostruire una voce umana** in modo estremamente realistico riproducendo caratteristiche come timbro, intonazione, inflessioni e persino piccoli dettagli come ritmo del respiro o pause del parlato. L'elaborazione sempre più avanzata di questi elementi permette ai cybercriminali di **creare copie vocali** di qualità sorprendente, rendendo le **imitazioni** sempre più **difficili da distinguere** dalla voce originale.



Mantieni sempre un **atteggiamento prudente** quando ricevi richieste inattese o urgenti e, alla richiesta di denaro, codici o informazioni sensibili, **interrompi la conversazione e verifica l'identità della persona** richiamandola tramite un numero ufficiale o già salvato nei contatti.

4 CASI REALI DI INCIDENTI DOVUTI AD UN USO MALEVOLO DELL'IA

Attacco deepfake ad una multinazionale di Hong Kong

Nel 2024, una nota società di ingegneria britannica con sede ad **Hong Kong** è stata vittima di uno dei più sofisticati attacchi informatici basati su **deepfake** mai registrati. Tutto è iniziato con un'e-mail apparentemente inviata dall'ufficio britannico dell'azienda, che invitava un dipendente a partecipare a una videoconferenza urgente per autorizzare una transazione. Durante la riunione *online*, il dipendente ha interagito con quello che sembrava essere il dirigente e alcuni colleghi: **volti, voci e movimenti erano perfettamente realistici**. In realtà, nessuna delle persone presenti era autentica, si trattava di **repliche digitali generate con l'Intelligenza Artificiale**. Convinto della legittimità della richiesta, il dipendente ha autorizzato un **bonifico di 25,6 milioni di dollari** verso un conto controllato dai truffatori.

Violazione di dati in un sistema scolastico basato su IA

Negli Stati Uniti, molti distretti scolastici stanno adottando sistemi di **sorveglianza digitale alimentati dall'Intelligenza Artificiale**, progettati per analizzare le attività online degli studenti e individuare segnali di comportamenti pericolosi, come autolesionismo, bullismo o minacce di violenza.

Tuttavia, durante un'indagine, è emerso che circa **3.500 documenti riservati**, compresi *report* su salute mentale, note disciplinari e conversazioni private, **erano accessibili pubblicamente online**, non protetti da alcun *firewall* o meccanismo di autenticazione.

L'incidente ha sollevato **preoccupazioni profonde tra genitori ed esperti di privacy**, che hanno messo in discussione la legittimità di questi sistemi, alimentando un ampio dibattito nazionale sul ruolo dell'Intelligenza Artificiale nelle scuole.





Intervento realizzato dal Ministero dell'Istruzione e del Merito e finanziato a valere sul Fondo per la gestione della cybersicurezza nell'ambito della misura 71 e 73 del Piano di Implementazione della Strategia Nazionale di Cybersicurezza 2022-2026 dell'Agenzia per la Cybersicurezza Nazionale



www.unica.istruzione.gov.it