AI come scudo: applicazioni e sfide nell'uso della Generative AI per contrastare le minacce cyber

Vincenzo
Calabrò

Funzionario alla Sicurezza CIS - Ministero dell'Interno
Professore a.c. di Tecnologie per la Sicurezza Informatica

Security4Business

4ª edizione 2025

4 Marzo 2025
Milano

IKN ITALY®
YOUR KNOWLEDGE NETWORK

Agenda

- opportunità della GenAI per migliorare la cyber resilience
- use cases di applicazioni GenAI nel cyber security lifecycle
- rischi di sicurezza specifici della GenAI
- rischi correlati all'utilizzo della GenAI nei processi aziendali

La GenAI può aiutare a migliorare la cyber resilience?

Fattori critici di *insuccesso* per la cyber resilience

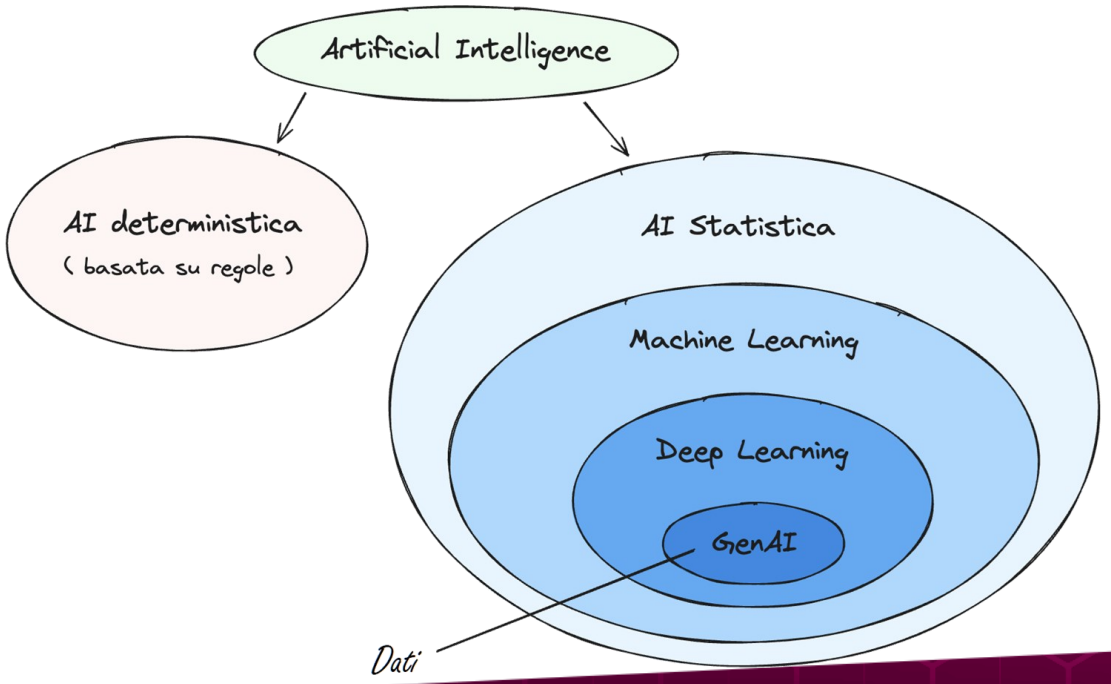
- **Intensa digitalizzazione** degli asset aziendali *(critici e non)*
- **Proliferazione dei dispositivi** (desktop, mobile, embedded, OT, IIOT)
- **Dipendenza dai fornitori** di servizi o soluzioni digitali
- **Incremento della complessità/fragilità** dei sistemi e dei dati
- **Estensione della superfice** di attacco *(edge e cloud)*
- **Aumento delle minacce** *(in termini quantitativi e qualitativi)*
- **Sofisticazione degli attacchi** *(in termini quantitativi e qualitativi)*
- **Crescita della quantità dei dati** da analizzare per la security
- **Incombenza derivante dalla compliance** normativa e regolatoria
- **Carenza di esperti** in cybersecurity *(non stressati)*

Fattori critici di *successo* della GenAI applicata alla cybersecurity?

- **Identificare gli errori** umani nella configurazione dei sistemi
- **Aumentare l'efficienza** delle attività ripetitive di cybersecurity
- **Limitare e gestire gli avvisi** di cybersecurity
- **Ridurre i tempi** di analisi e risposta alle minacce
- **Identificare e prevedere le nuove minacce** alla sicurezza informatica
- **Gestire la capacità** del personale nella cybersecurity

Qual è il contributo concreto della GenAI?

L'intelligenza artificiale generativa (GenAI) è un sottoinsieme dell'intelligenza artificiale (AI) che impiega **l'apprendimento automatico** e le **reti neurali profonde** per analizzare **grandi serie di dati e** creare **risultati simili, ma nuovi, simulando il ragionamento umano**.



Quando alimentiamo un modello di GenAI con **i dati** di addestramento, apprende gli **schemi**, le **strutture** e le **relazioni** sottostanti e crea una **rappresentazione** compressa **dei dati** in uno spazio ad alta dimensionalità.

Queste informazioni vengono poi elaborate per **generare nuovi risultati attraverso i modelli di GenAI**, quali:

- **Variational Autoencoders (VAE)**
- **Generative Adversarial Networks (GAN)**
- **Diffusion Model**
- **Transformer**

In pratica, la GenAI può assistere la cybersecurity a:

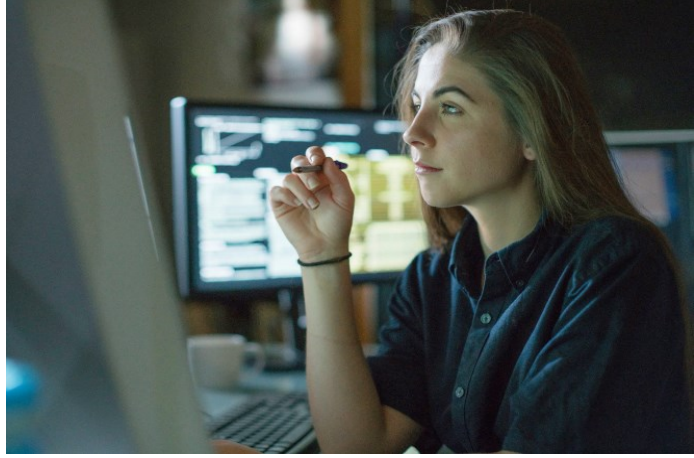
- **Ridurre il rumore dei dati**
- **Rilevare anomalie**
- **Avvisare dei rischi sui dati**
- **Simulare attacchi**

Use case per la cybersecurity

1. Previsione e prevenzione delle minacce

Task	Ruolo della GenAI
Valutazione delle minacce e delle vulnerabilità	• Aggregare, correlare e analizzare i dati sulle minacce per identificare i rischi emergenti • Automatizzare la scansione delle vulnerabilità e dare priorità ai punti deboli ad alto rischio • Prevedere i percorsi di attacco futuri e dare priorità alle azioni di mitigazione
Ottimizzazione della postura di sicurezza	• Generare raccomandazioni attuabili per migliorare la sicurezza • Automatizzare l'assegnazione dei compiti, tracciare i progressi e segnalare i ritardi • Automatizzare le patch di sicurezza, gli aggiornamenti della configurazione e l'hardening del sistema
Convalida e miglioramento continuo	• Analizzare l'impatto delle modifiche prima dell'implementazione per ridurre le interruzioni • Fornire avvisi e rapporti in tempo reale sulle minacce emergenti e sulle best practice • Monitoraggio continuo della sicurezza, generazione di report e monitoraggio dei progressi

Use case per la cybersecurity



2. Gestione delle minacce in tempo reale

Task	Ruolo della GenAI
Rilevamento e analisi di incidenti	• Rilevare anomalie e schemi sospetti in tempo reale • Distinguere le attività dannose dalle anomalie benigne • Dare priorità agli incidenti in base alla gravità e al potenziale impatto sull'azienda
Contenimento e eradicazione	• Automatizzare le azioni di contenimento: isolare i sistemi, bloccare gli IP, mettere in quarantena i file • Analizzare il comportamento del malware e consigliare un contenimento mirato • Individuare i punti deboli del sistema e fornire indicazioni per la correzione
Recupero e ripristino	• Assistere nel recupero dei dati e nel ripristino dei sistemi agli stati precedenti all'incidente • Monitoraggio della reinfezione o delle minacce persistenti dopo la bonifica
Apprendimento e miglioramento continuo	• Affinare i modelli di minaccia e migliorare il rilevamento in base alle nuove minacce • Consigliare aggiornamenti alle policy di sicurezza e alle misure di hardening • Automatizzare le attività, semplificate la risposta agli incidenti e integrare i suggerimenti dell'intelligenza artificiale

Use case per la cybersecurity

3. Attività post-incidente

Task	Ruolo della GenAI
Contenimento delle minacce	• Raccomandare strategie di contenimento ottimali per limitare la diffusione degli attacchi • Identificare e disattivare gli account compromessi, bloccare gli IP sospetti e applicare i controlli di accesso • Fornire informazioni sulle capacità delle minacce informatiche, sui vettori di attacco e sugli indicatori di compromissione (IOC)
Mitigazione e recupero della funzionalità	• Scoprire i percorsi di attacco, le lacune di sicurezza e acquisire conoscenze sulle tecniche • Suggerire la correzione delle vulnerabilità, rafforzare le configurazioni e migliorare le politiche di sicurezza • Assistere nel recupero dei dati, nel ripristino del sistema e nella convalida dell'integrità dei dati
Revisione e miglioramento post-incidente	• Generare rapporti sugli incidenti, comprese le tempistiche, l'analisi degli attacchi e le azioni intraprese • Identificare le aree di miglioramento nel rilevamento, nel contenimento e nei processi • Migliorare i modelli delle minacce, i controlli di sicurezza e i playbook di risposta agli incidenti sulla base dell'analisi post-incidente

Use case per la cybersecurity

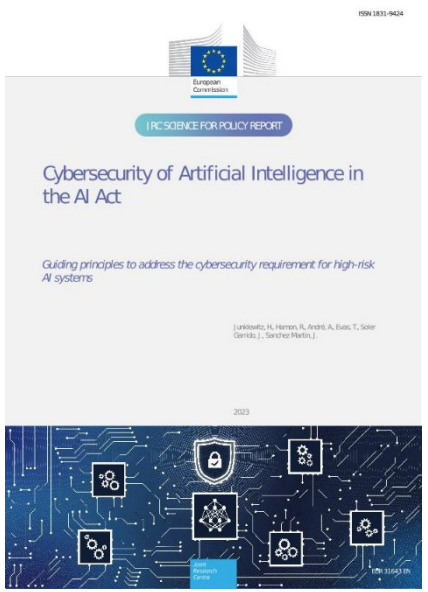
4. Miglioramenti e adattamenti continui

Task	Ruolo della GenAI
Raccolta e analisi dei dati	• Aggregare e analizzare i dati provenienti da varie fonti per identificare le tendenze e i modelli • Automatizzare le valutazioni della sicurezza ed evidenziare le aree da migliorare • Generare simulazioni di phishing e monitorare le percentuali di clic e i comportamenti di segnalazione
Evoluzione e automazione della strategia	• Identificare i vettori di attacco emergenti e prevedere le minacce in base all'analisi delle tendenze • Affinare i modelli di minaccia e migliorare l'accuratezza del rilevamento delle minacce in evoluzione • Consigliare aggiornamenti dei criteri, modifiche alla configurazione e miglioramenti dei controlli di sicurezza • Semplificare le operazioni con l'automazione e l'orchestrazione intelligente
Valutazione e miglioramento continuo	• Tracciare le metriche (ad esempio, tempo di rilevamento, tempo di riparazione, numero di incidenti) e fornire informazioni utili all'improvement delle performance • Valutare e perfezionare continuamente i processi automatizzati per ottenere prestazioni ottimali • Fornire avvisi in tempo reale e informazioni utili al personale addetto alla sicurezza

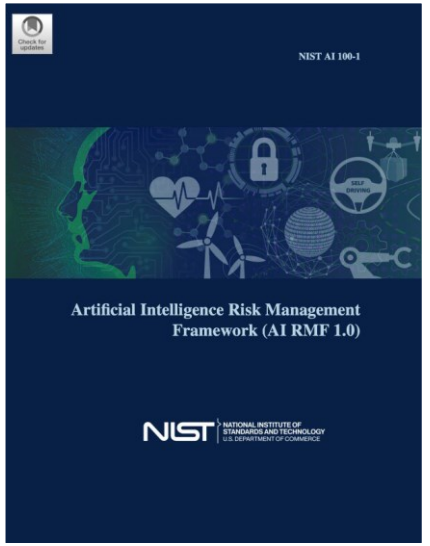
Quant'è «*safe*» la GenAI per la cybersecurity?

L'output di un modello di GenAI è **una previsione statistica** generata utilizzando i dati forniti per l'addestramento; pertanto, occorre che **i dati per il training del modello siano quantitativamente e qualitativamente rilevanti**.
Preferire soluzioni di cybersecurity con GenAI che indichino i modelli e i dati utilizzati per l'addestramento

Inoltre, i modelli di GenAI possono contenere ulteriori **criticità** intrinseche al modello stesso.

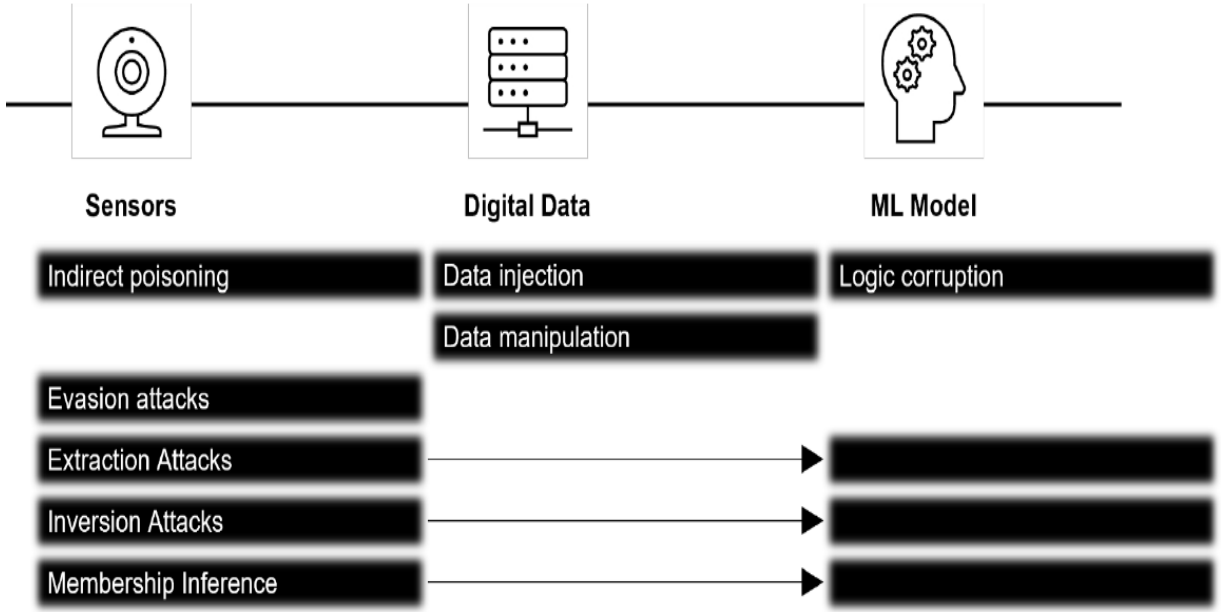


Criticità	Contromisura
Allineamento	Meccanismi di feedback umani per perfezionare l'output oppure Sistemi di correzioni automatica
Solidità	Monitoraggio continuo del modello e dell'ambiente Applicazione costante degli aggiornamenti del modello
Trasparenza	Raccogliere un feedback umano sull'interpretabilità
Privacy/Breach	Effettuare penetration testing Anonimizzazione dei dati
Bias/Pregiudizio	Monitoraggio continuo delle prestazioni del modello
Compliance	Recepire le prescrizioni normative e regolatorie



La GenAI è vulnerabile?

L'intelligenza artificiale generativa è integrata in sistemi IT/OT; pertanto, è esposta alle **cyber minacce** e alle **minacce specifiche dei sistemi GenAI**. Questi ultimi possono colpire **obiettivi** diversi e in **momenti** diversi. Occorre includere queste minacce nel **risk assesment.**



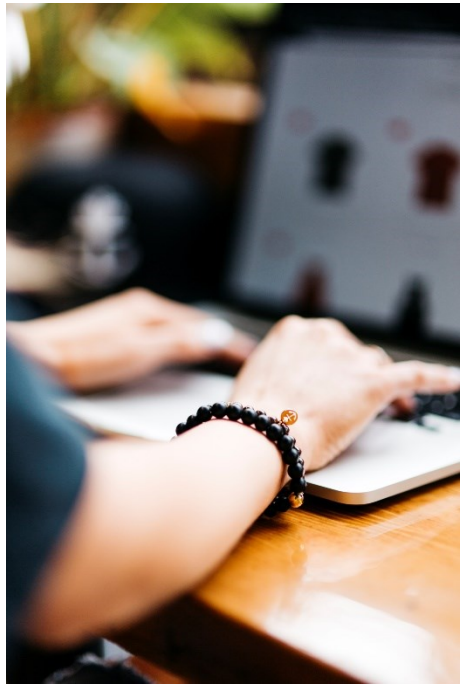
Minacce specifiche per GenAI

MITRE ATLAS™

Adversarial Threat Landscape for Artificial-Intelligence Systems

Contromisure specifiche per GenAI

Dov’è utilizzata la GenAI in azienda?



Dispositivi end user

Integrata nei servizi IT / OT

Inserita nei servizi di security

Violazioni della privacy, del copyright o del segreto industriale

Shadow AI !!!

Regolamento all’uso corretto della GenAI
Alimentare la GenAI con dati proprietari

Security4Business

4ª edizione 2025

4 Marzo 2025
Milano

IKN ITALY®
YOUR KNOWLEDGE NETWORK

GRAZIE!

www.vincenzocalabro.it