Cybersecurity & AI

eGeH ‘26

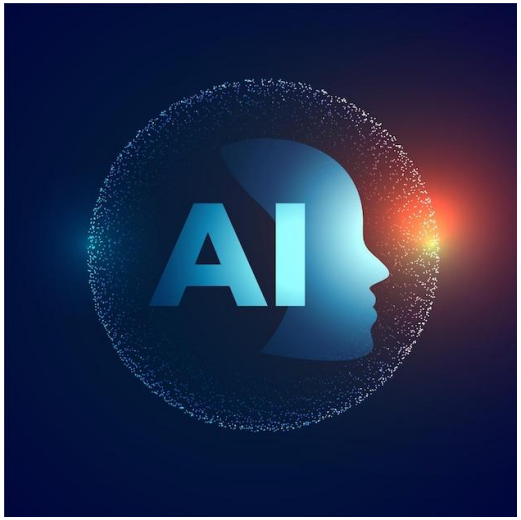
Desio 10 Luglio 2026

Andrea Capra

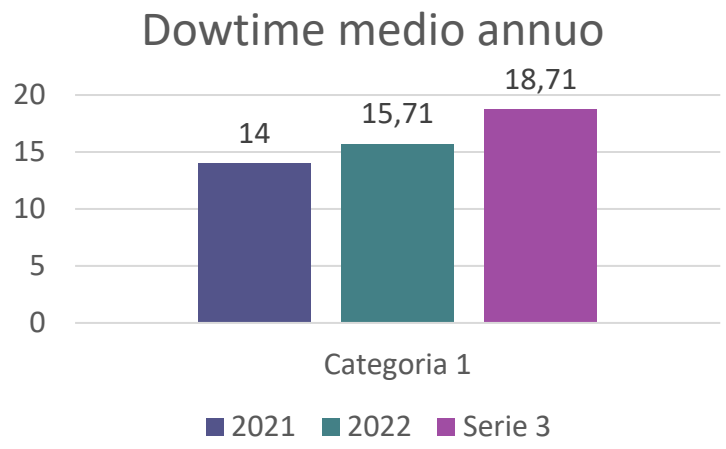
CASPO Consulting

IITM

AITIM
INFORMATICA MEDICA



Impatto dei soli attacchi ransomware (2021-2023)



Fonte: studio Comparitech citato da CyberWire (2023)

Questi valori rappresentano **interruzioni dei sistemi clinici e amministrativi**, incluse:

- EHR/Cartelle cliniche elettroniche
- Sistemi di laboratorio e radiologia
- Sistemi di prenotazione e accettazione
- Portali pazienti
- Infrastrutture di billing e assicurazione

Il downtime non significa necessariamente “ospedale fermo”, ma:

- **Ritorno a carta e penna**
- **Rinvio di interventi non urgenti**
- **Blocchi nei sistemi diagnostici**
- **Impossibilità di accedere ai dati clinici**

Totale downtime nel settore sanitario USA (dal 2016 al 2023) (dati 2024 e 2025 non ancora disponibili)
6.347 giorni complessivi → 17,4 anni di indisponibilità cumulata

Non ci sono dati ufficaili sull’idisponibilita dei sistemi per l’italia
Casi ASL1 Abruzzo e Multimedica (2023) – Disservizio per i pazienti di oltre 15 gg

A questi disservizi ovviamente si aggiunge spesso la esfilatrazione e diffusione non autorizzata dei dati sensibili dei pazienti

L’AI ridefinisce la natura degli attacchi informatici
…… e ha già cominciato a sostituire gli hacker

Da strumento di supporto a delega operativa lungo la catena d’attacco

IERI · Strumento di supporto

- Accelerare la scrittura di codice
- Raccogliere informazioni
- Analizzare le vulnerabilità

OGGI · Delega operativa

- Il ruolo dei modelli cambia qualitativamente
- L’AI non assiste più l’attaccante
- Inizia a sostituirlo in alcune fasi della catena d’attacco

L’intelligenza artificiale non accelera più soltanto gli attacchi informatici: inizia a eseguirne alcune fasi in autonomia. Dai modelli linguistici agli AI agent, cambia la cyber kill chain e diventa necessario ripensare difesa, governance e resilienza digitale. L’AI non assiste più l’attaccante: inizia a sostituirlo in alcune fasi della catena d’attacco.

Dall’uso “strumentale” dell’AI al cyber spionaggio assistito

Già nel 2025 Anthropic aveva pubblicato uno dei primi avvisi strutturati sull’uso offensivo dei propri modelli da parte di attori state-sponsored**:** campagne attribuite a gruppi legati alla Cina che sfruttavano il modello Claude per supportare attività di cyber spionaggio e intelligence.

AI non ancora autonoma → **Ruolo di acceleratore cognitivo**: analisi di target, traduzione di contenuti, generazione di script e supporto alla pianificazione.

Mancava la componente strategica dell’attacco.

↓ Elemento cruciale

Riduzione drastica della soglia di competenza necessaria per condurre operazioni sofisticate.

Ransomware, il salto qualitativo: l'emergere degli AI agent offensivi

Fase successiva

AI agent capaci non solo di generare contenuti ma di interagire con ambienti esterni, API, sistemi operativi e infrastrutture cloud

Capacità di eseguire sequenze operative complesse in modo semi-autonomo e adattativo, "ragionando" sul risultato delle proprie azioni.

Operazione JADEPUFFER

analizzata dall'azienda di sicurezza Sysdig

prima campagna ransomware end-to-end guidata da un modello linguistico

L'agente AI ha compromesso un server vulnerabile, raccolto credenziali cloud, esplorato infrastrutture interne, compromesso servizi secondari, ottenuto privilegi elevati e infine cifrato e distrutto dati, senza intervento umano diretto.

l'agente è stato stato in grado di correggere autonomamente i propri errori operativi e completare la correzione in più fasi in 31 secondi

Come l'AI può cambiare un attacco ransomware

Storicamente, un attacco ransomware richiedeva competenze specialistiche e una netta separazione tra fasi: accesso iniziale, escalation, movimento laterale, esfiltrazione e cifratura. Ogni fase implicava strumenti diversi e operatori distinti.

Con gli AI agent questa separazione si riduce drasticamente.

Un singolo sistema può oggi individuare vulnerabilità pubbliche sfruttabili (CVE), sfruttarle automaticamente, raccogliere credenziali e token cloud, esplorare ambienti interni e identificare asset critici. Può esfiltrare dati, eseguire cifratura o distruzione delle informazioni e, infine, generare note di riscatto e gestire le comunicazioni legate all'estorsione.

La cyber kill chain passa da un processo multi-attore a un flusso continuo automatizzato.

Con i nuovi Modelli agentici (Mythos, Fable, ecc.) la AI diventa un moltiplicatore di potenza offensiva sul fronte geopolitico :

- riduce il tempo necessario per operazioni di intrusione;
- abbassa la soglia di competenza tecnica;
- aumenta la scala delle campagne;
- automatizza il ciclo di attacco completo.

In questo contesto, la disponibilità di modelli meno vincolati da restrizioni etiche o normative può rappresentare un vantaggio strategico significativo per attori statali o para-statali.

Dalla sicurezza reattiva alla sicurezza adattiva

I modelli difensivi tradizionali non sono più adatti

La sicurezza basata su signature, IOC statici o regole predefinite è sempre meno efficace contro attori che non seguono più pattern deterministici.
Gli AI agent possono modificare il proprio comportamento in tempo reale, adattandosi alle difese che incontrano lungo il percorso. Questo significa che non seguono più schemi fissi, ma possono generare continuamente varianti diverse delle stesse tecniche di attacco, rendendo più difficile il riconoscimento basato su firme o pattern statici. Inoltre, sono in grado di simulare il comportamento umano, replicando tempi, interazioni e modalità operative credibili per eludere i sistemi di detection.

Necessario un cambiamento di paradigma nella sicurezza

Necessari modelli di detection comportamentale avanzata (User and Entity Behavior Analytics, UEBA), capaci di analizzare ciò che un utente o un sistema fa nel tempo e non solo ciò che è.

Serve, inoltre, una capacità di analisi delle anomalie in tempo reale, un controllo molto più rigoroso delle identità secondo principi di Zero Trust esteso e l'isolamento dei workload legati all'AI e dei sistemi CI/CD.

Necessità di un monitoraggio continuo delle catene di supply chain software, oggi uno dei punti più esposti agli attacchi automatizzati.

Problemi di continuità operativa

È verosimile che **un presidio ospedaliero universitario non abbia definito le procedure volte a garantire la continuità operativa della propria struttura ?**
È possibile che lo scenario "indisponibilità del sistema informativo" non sia stato contemplato?

Sul tema della continuità operativa, lo schema di riferimento è la norma internazionale **ISO 22301:2019 (BCMS – Business Continuity Management Systems)**, finalizzata a definire "*la struttura e i requisiti per l'implementazione e il mantenimento di un sistema di gestione della continuità aziendale (BCMS) che sviluppi una continuità operativa adeguata alla quantità e al tipo di impatto che l'organizzazione può o non può accettare a seguito di un'interruzione.*"

Quando si parla di continuità operativa di un servizio (Service Continuity) è importante fare attenzione e non confonderla con la disponibilità del servizio (Service Availability), che invece viene definita in sede di progetto del sistema in funzione dei requisiti del servizio.

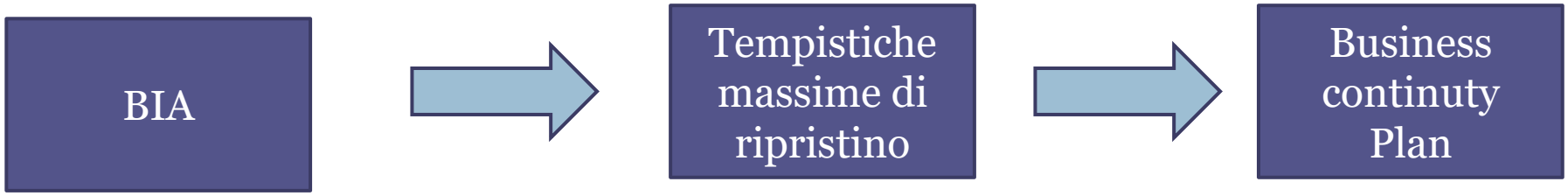
BIA – Business impact analisys

ISO/TS 22317:2015 (Guidelines for business impact analysis (BIA)).

Consente di fare delle valutazioni in termini di impatto sul business (business che, nel caso di un presidio ospedaliero, è l'erogazione dei servizi sanitari), considerando i vari fattori che possono intervenire in caso di interruzione del servizio.

Si prendono in considerazione i principali elementi che sono coinvolti nel processo di business, da fattori reputazionali a fattori economici, dalle possibili conseguenze sulla salute delle persone a considerazioni sugli aspetti normativi coinvolti. L'impatto, varia al passare del tempo, in genere in modo progressivo. **Più passa il tempo, maggiore sarà l'impatto complessivo sul business.**

Con questo strumento definiamo, quindi, in modo strutturato le tempistiche (**RTO – Recovery Time Objective**) entro le quali è opportuno (necessario) intervenire per rispristinare il servizio interrotto, affinché questa interruzione non determini degli impatti rilevanti sul business, e come questo impatto varia al passare del tempo.



Business Continuity Plan (BCP) - ossia il Piano di Continuità Operativa del servizio, viene predeterminato dall'organizzazione a fronte di **determinati scenari**, e attivato qualora un incident causi l'interruzione di un servizio (da non confondere con il piano di ripristino dell'operatività del sistema informativo che potrebbe ancora essere compromesso)..

Identificare preventivamente i possibili scenari

L'identificazione preventiva dei possibili scenari è funzionale all'identificazione di quelli che potrebbero essere i fattori determinanti per la compromissione del servizio (del tutto o in parte).
Per specifici contesti critici, come quelli dell'emergenza sanitaria, il parametro RTO dovrebbe essere "Near Zero" (ossia prossimo allo Zero), e il BCP deve essere rapidamente applicabile.

Il **triage di un paziente** in arrivo al pronto soccorso viene svolto dall'infermiere, ed in funzione del codice di emergenza attribuito al paziente, si determina la priorità di accesso e di cura.
In mancanza del sistema informativo, le informazioni necessarie si annotano in modalità cartacea e l'informazione segue il paziente nel processo di gestione all'interno dei reparti.

Durante chiamata al **numero di emergenza 112**, se il sistema informativo che gestisce la comunicazione tra il chiamante e i **PSAP di livello 2 (Public Safety Answering Point)** dovesse risultare improvvisamente indisponibile, l'operatore continuerebbe a **gestire la chiamata di emergenza su carta, e comunicherebbe le informazioni per la gestione delle chiamate via telefonica** e non attraverso il sistema informativo attraverso la scheda contatto, come avviene in condizioni normali.

Nel frattempo l'attuazione del **DRP (Disaster Recovery Plan)** consentirà di ripristinare, dal punto di vista tecnico, il sistema informativo compromesso.