Threat Modeling nel 2026: analisi approfondita, aggiornata e operativa

32% degli attacchi parte da una vulnerabilità sfruttata. 11% da una telefonata. 9% da credenziali rubate. I numeri del report M-Trends 2026 di Mandiant raccontano una storia precisa: chi attacca non forza più le porte, le apre con le chiavi giuste, al momento giusto, spesso senza che nessun alert si attivi.

In questo scenario, sapere dove si è vulnerabili non basta più. Serve sapere come un avversario reale si muoverebbe attraverso i propri sistemi, quali identità sfrutterebbe, quali agenti AI manipolerebbe, quali fornitori userebbe come testa di ponte. Questa capacità ha un nome consolidato, threat modeling, e nel 2026 ha smesso di essere una pratica da specialisti per diventare un requisito operativo, regolamentare e strategico per qualsiasi organizzazione che gestisca infrastrutture digitali critiche.

Questa analisi percorre lo stato dell'arte della disciplina: dai framework classici in evoluzione (STRIDE, PASTA, LINDDUN) ai nuovi standard AI-native come MITRE ATLAS v5.4.0, MAESTRO e l'OWASP Top 10 per le applicazioni agentiche, fino a un percorso di maturità praticabile anche per chi parte da zero.

Threat Modeling: una disciplina che non si può più rimandare

Per anni il threat modeling è rimasto confinato ai margini del ciclo di sviluppo software: un esercizio formale, spesso delegato a un singolo specialista di sicurezza, raramente aggiornato dopo il rilascio in produzione. Nel 2026 quella

stagione è definitivamente chiusa.

Tre forze convergenti hanno trasformato il threat modeling da pratica opzionale a imperativo strategico. La prima è un threat landscape che si industrializza a velocità senza precedenti: gli attaccanti operano come organizzazioni strutturate, con divisione del lavoro, specializzazione e catene di fornitura del crimine che comprimono i tempi di compromissione a pochi secondi. La seconda è la diffusione pervasiva di sistemi di intelligenza artificiale agentici, che ridisegnano la superficie d'attacco in modo radicalmente diverso da qualsiasi tecnologia precedente. La terza è un quadro regolamentare europeo – AI Act, NIS2, Cyber Resilience Act, DORA – che impone la formalizzazione del risk modeling come requisito di conformità verificabile, non più come buona pratica volontaria.

Questa analisi ricostruisce lo stato dell'arte del threat modeling nel 2026 in modo operativo: framework in evoluzione, domini applicativi prioritari, strumenti disponibili, errori comuni da evitare e un percorso di maturità praticabile anche per le organizzazioni che partono da zero.

1. Il contesto: un threat landscape che si industrializza

Il punto di partenza non può che essere il dato più rilevante dell'anno. Il report M-Trends 2026 di Mandiant – basato su centinaia di interventi di incident response condotti a livello globale nel 2025 – fotografa un panorama in cui i cyberattacchi sono diventati più veloci, più coordinati e sempre più professionalizzati. Gli attaccanti riescono a trasferire accessi tra diversi soggetti in meno di trenta secondi, comprimendo drasticamente le finestre di risposta per i difensori e rendendo obsoleti i playbook tradizionali.

Il global median dwell time – cioè il tempo medio che un

attaccante trascorre non rilevato all'interno di un sistema compromesso – è salito a quattordici giorni, in crescita rispetto agli undici dell'anno precedente. L'aumento è trainato in larga misura da attività di spionaggio a lungo termine e da operazioni di IT worker collegati alla Corea del Nord, che infiltrano organizzazioni tecnologiche occidentali attraverso false identità professionali.

Sul fronte dei vettori iniziali di compromissione, lo sfruttamento di vulnerabilità (exploit) si conferma in cima alla classifica con il 32% dei casi, seguito dal voice phishing all'11%, dal prior compromise al 10% e dalle credenziali rubate al 9%. Il web compromise contribuisce per l'8%, mentre email phishing e insider threat si attestano entrambi al 6%. Questi numeri hanno implicazioni dirette sul threat modeling: un modello che prioritizza esclusivamente le vulnerabilità tecniche sottostima gravemente i vettori di compromissione basati sull'identità e sull'ingegneria sociale.

Un dato strutturale da tenere a mente riguarda il cambiamento di paradigma degli attaccanti. Il panorama delle minacce si è spostato verso un approccio "log in rather than break in": gli avversari sfruttano credenziali valide, session token e accessi federati per aggirare le difese perimetrali tradizionali senza mai attivare un singolo alert di rilevamento perimetrale. Questo sposta il baricentro del threat modeling dall'analisi delle vulnerabilità tecniche alla modellazione dell'abuso di identità e degli accessi legittimi.

Sul fronte dell'intelligenza artificiale, il Google Threat Intelligence Group conferma che gli avversari stanno integrando l'AI per accelerare il ciclo d'attacco. Famiglie di malware come PROMPTFLUX e PROMPTSTEAL interrogano attivamente modelli linguistici durante l'esecuzione per eludere il rilevamento; le cosiddette "distillation attack" minacciano la proprietà intellettuale estraendo la logica proprietaria e i dati di addestramento specializzati di modelli ad alto valore. Al tempo stesso, il credential stealer QUIETVAULT è stato

osservato mentre analizzava le macchine target alla ricerca di strumenti AI da riga di comando, eseguendo prompt predefiniti per individuare file di configurazione.

Un avvertimento prezioso per chi è tentato di sovrastimare la minaccia AI: la grande maggioranza delle intrusioni di successo continua a derivare da fallimenti umani e sistemici fondamentali – patch management carente, MFA assente, privilege sprawl non governato – e non da attacchi AI sofisticati. Il threat modeling efficace non deve inseguire la novità a scapito dei fondamentali.

2. Che cos'è il threat modeling e perché va reinterpretato nel 2026

Il threat modeling è, nella sua definizione più solida, una disciplina strutturata per identificare, analizzare e comunicare le minacce nel contesto della protezione di un asset di valore. Applicato al software, abilita decisioni informate sui rischi di sicurezza applicativa e produce – oltre al modello stesso – un elenco prioritizzato di miglioramenti da applicare a concept, requisiti, progettazione o implementazione.

Il Threat Modeling Manifesto, redatto nel 2020 da un gruppo di practitioner internazionali, ha codificato quattro domande fondamentali che strutturano qualsiasi sessione di threat modeling: su cosa stiamo lavorando? Cosa può andare storto? Cosa possiamo fare al riguardo? Il lavoro svolto è stato efficace? Queste domande, pur nella loro apparente semplicità, sono ancora oggi la bussola più affidabile per orientarsi in un processo che rischia di diventare esercizio burocratico invece di strumento operativo.

Il thread modeling tradizionale si è evoluto attorno al software deterministico: percorsi di codice noti, input e output prevedibili, modalità di fallimento relativamente

stabili. I sistemi di intelligenza artificiale – in particolare quelli generativi e agentici – rompono molte di queste assunzioni. Un sistema AI è probabilistico: lo stesso input può produrre output diversi in esecuzioni successive. Tratta la conversazione e le istruzioni come parte di un singolo flusso di input in cui il testo avversariale può essere interpretato come intento eseguibile. Può invocare API esterne, persistere stato in memoria e attivare workflow autonomamente, con effetti a cascata su componenti non direttamente modellati.

Il 2026 impone quindi una doppia lettura del threat modeling: la consolidazione delle metodologie classiche in pipeline DevSecOps mature, e la costruzione di framework radicalmente nuovi capaci di modellare sistemi probabilistici, non deterministici, con autonomia d'azione e memoria persistente.

3. Framework metodologici classici: evoluzione e stato dell'arte

STRIDE

Creato da Microsoft, STRIDE rimane il framework più diffuso e pedagogicamente più accessibile. L'acronimo copre sei categorie di minaccia: Spoofing (impersonificazione di un'entità legittima), Tampering (manomissione di dati o codice), Repudiation (ripudio di un'azione), Information Disclosure (divulgazione non autorizzata di informazioni), Denial of Service (saturazione di risorse) ed Elevation of Privilege (acquisizione di permessi superiori a quelli legittimi).

STRIDE si applica in modo naturale insieme ai Data Flow Diagram, analizzando ogni elemento del diagramma – processi, archivi dati, flussi, entità esterne – contro le sei categorie. Nel 2026 viene estensivamente usato nelle pipeline DevSecOps integrate, ma richiede estensioni esplicite per coprire le superfici AI: la categoria "Spoofing" va estesa

all’impersonificazione tramite prompt injection, “Tampering” al data poisoning nei dataset di addestramento, “Elevation of Privilege” all’escalation ottenuta manipolando il contesto di un agente autonomo.

PASTA

PASTA (Process for Attack Simulation and Threat Analysis) è una metodologia attacker-centric articolata in sette fasi: definizione degli obiettivi di business, definizione dello scope tecnico, decomposizione dell’applicazione, analisi delle minacce, analisi delle vulnerabilità, simulazione degli attacchi, analisi del rischio e impatto. È particolarmente indicata per threat modeling di sistemi ad alta complessità regolamentare, poiché allinea l’analisi delle minacce agli obiettivi di business e all’impatto operativo misurabile. Una banca, un’azienda sanitaria o un operatore di infrastrutture critiche troverà in PASTA la metodologia più aderente al proprio modello di governance del rischio.

LINDDUN

LINDDUN è una metodologia focalizzata sulla privacy, progettata per identificare minacce relative a Linkability (possibilità di collegare dati appartenenti allo stesso soggetto), Identifiability (possibilità di identificare un individuo da dati aggregati), Non-repudiation, Detectability, Disclosure of Information, Unawareness (mancanza di consapevolezza dell’interessato) e Non-compliance. Supporta nativamente GDPR, HIPAA e normative analoghe. Nel contesto europeo del 2026 – con l’AI Act in piena applicazione – LINDDUN sta acquisendo rilevanza crescente per le organizzazioni che devono modellare i rischi dei sistemi AI sull’identità e sulla protezione dei dati personali degli utenti.

Attack Trees

Uno degli approcci più longevi, ancora ampiamente usato per

analizzare scenari d’attacco complessi in modo gerarchico. La radice dell’albero è l’obiettivo dell’attaccante; ogni nodo figlio rappresenta un modo per raggiungere il nodo genitore. Nel 2026 gli attack tree vengono spesso generati automaticamente da strumenti AI-assisted come IriusRisk o Microsoft Threat Modeling Tool, riducendo il tempo necessario per coprire scenari complessi.

Framework ibridi

Gli approcci più maturi combinano più metodologie per ottenere copertura più ampia. Una configurazione comune nel 2026 è l’uso di STRIDE per la struttura di base, LINDDUN per la dimensione privacy, e PASTA per la valutazione dell’impatto business, il tutto mappato sulle tattiche e tecniche di MITRE ATT&CK per verificare la copertura rispetto al threat landscape reale.

4. ## La tabella comparativa: scegliere il framework giusto



5. ## I nuovi framework AI-native: la svolta strutturale del 2026

Questa è la frontiera più rilevante e meno consolidata del threat modeling nel 2026. L’insufficienza dei framework classici per i sistemi AI non è un’opinione: è un dato strutturale riconosciuto da MITRE, OWASP, CSA e Microsoft con aggiornamenti formali dei rispettivi framework.

MITRE ATLAS v5.4.0

MITRE ATLAS (Adversarial Threat Landscape for Artificial-Intelligence Systems) è la base di conoscenza di riferimento per le minacce ai sistemi AI, organizzata come MITRE ATT&CK ma specificamente progettata per modelli ML e sistemi AI in

produzione. L’aggiornamento di novembre 2025 ha portato il framework a sedici tattiche, ottantaquattro tecniche, trentadue mitigazioni e quarantadue case study documentati. L’aggiornamento successivo del febbraio 2026 (v5.4.0) ha aggiunto ulteriori tecniche incentrate sugli agenti AI, tra cui “Publish Poisoned AI Agent Tool” ed “Escape to Host”.

L’aggiornamento 2026 di MITRE ATLAS sposta il focus dagli attacchi centrati sul modello all’esposizione al livello di execution layer: il threat modeling deve ora tenere conto del chaining autonomo di workflow, della persistenza dell’autorità delegata e del rischio di orchestrazione a livello API. I team di sicurezza non possono più limitare le valutazioni agli input e agli output del modello: devono valutare come i sistemi AI interagiscono con l’infrastruttura enterprise nel tempo, attraverso tool call, accessi a memoria persistente e deleghe di autorità.

Un caso documentato particolarmente istruttivo è SesameOp, aggiunto ad ATLAS a fine 2025: una backdoor che sfrutta le API dei servizi di agenti AI come canale di command and control, mascherando attività malevole nel normale traffico verso i provider AI. Un altro caso documentato riguarda il Financial Transaction Hijacking tramite Microsoft 365 Copilot, che dimostra come un assistente AI aziendale possa essere sfruttato per operazioni finanziarie non autorizzate attraverso la manipolazione del suo contesto.

MAESTRO

MAESTRO (Multi-Agent Environment, Security, Threat, Risk, and Outcome) è il framework della Cloud Security Alliance specificamente progettato per gli agenti AI autonomi. A differenza di ATLAS – che è essenzialmente una tassonomia delle minacce – MAESTRO è un framework di processo che guida l’analisi del rischio lungo l’intero ciclo di vita agentico, organizzata attorno agli elementi di sistema AI (ambiente, strumenti, memoria, identità, comunicazione inter-agente).

Include un tool AI-powered per il threat modeling che genera automaticamente scenari di rischio a partire dall'architettura del sistema.

OWASP Top 10 per le applicazioni agentiche (2026)

La prima edizione dell'OWASP Top 10 for Agentic Applications cataloga i dieci rischi più critici per i sistemi AI agentici: Agent Goal Hijacking (ASI01), Tool Misuse (ASI02), Identity & Privilege Abuse (ASI03), Supply Chain Vulnerabilities (ASI04), Unexpected Code Execution (ASI05), Memory & Context Poisoning (ASI06), Insecure Inter-Agent Communication (ASI07), Cascading Failures (ASI08), Human-Agent Trust Exploitation (ASI09) e Rogue Agents (ASI10). La lista è complementare — non sostitutiva — all'OWASP Top 10 per le applicazioni LLM, che mantiene la Prompt Injection al primo posto con l'aggiunta di System Prompt Leakage e Vector & Embedding Weaknesses.

6. Il tool landscape: gli strumenti disponibili nel 2026

Il threat modeling è efficace solo se viene praticato sistematicamente, e la sistematicità richiede strumenti adeguati. Il mercato nel 2026 si divide in tre categorie.

Strumenti open source e gratuiti. Microsoft Threat Modeling Tool rimane il punto di partenza più accessibile per chi usa STRIDE su applicazioni Windows e cloud Azure. OWASP Threat Dragon è l'alternativa open source multi-piattaforma, con supporto per DFD e modelli STRIDE esportabili. MITRE ATLAS Navigator e il relativo Arsenal permettono di mappare i propri controlli di sicurezza sulle tattiche ATLAS e di identificare le lacune di copertura specifiche per i sistemi AI.

Piattaforme commerciali integrate. IriusRisk è oggi la piattaforma più matura per il threat modeling enterprise integrato nel ciclo DevSecOps, con supporto per più framework, generazione automatica di threat model a partire dai diagrammi

architetturali, integrazione con Jira e sistemi di ticketing, e un motore basato su regole aggiornato continuamente. ThreatModeler offre funzionalità analoghe con enfasi sull'automazione tramite AI. SD Elements di Security Compass si distingue per la capacità di tradurre il threat model in requisiti di sicurezza specifici per sviluppatori, riducendo il gap tra modellazione e implementazione.

Strumenti AI-assisted di nuova generazione. Nel 2026 emergono soluzioni che usano modelli linguistici per generare threat model a partire da descrizioni testuali dell'architettura, diagrammi esistenti o documentazione di progetto. Questi strumenti accelerano significativamente la fase di elicitazione delle minacce, ma richiedono revisione esperta dell'output: i modelli linguistici tendono a produrre minacce plausibili ma non necessariamente prioritizzate rispetto al contesto specifico dell'organizzazione.

7. I domini applicativi prioritari nel 2026

7.1 Threat modeling dell'identità

La surface d'attacco è oggi identity-driven. Gli avversari scelgono di accedere anziché forzare, sfruttando credenziali valide, session token e accessi federati per aggirare le difese perimetrali. Il voice phishing – undici per cento dei vettori iniziali secondo M-Trends 2026 – ha soppiantato il phishing tradizionale via email per gli attacchi mirati di alto valore, perché la componente umana in tempo reale rende molto più difficile il rilevamento automatico.

Il threat modeling dell'identità deve coprire scenari che i framework tradizionali non modellano esplicitamente: abuso di SSO e identity provider federati, session hijacking post-autenticazione MFA, exploitation delle integrazioni SaaS (OAuth scope abuse), privilege escalation tramite service account e managed identity nei cloud provider, e social engineering degli helpdesk IT che possono essere indotti a

resettare credenziali o bypassare procedure di verifica.

7.2 Threat modeling dei sistemi AI agentici

Il threat modeling AI è intrinsecamente difficile per tre ragioni strutturali. Prima: il comportamento dei sistemi AI è probabilistico e dipende dal contesto, rendendo impossibile enumerare in anticipo tutti i percorsi di esecuzione. Seconda: la superficie d'attacco si estende attraverso tutto ciò che l'agente può fare — tool call, accessi a database, invio di email, scrittura di codice — non solo attraverso i suoi input diretti. Terza: gli effetti di una compromissione possono propagarsi rapidamente attraverso sistemi connessi: un singolo agente compromesso può avvelenare l'87% delle decisioni downstream entro quattro ore, secondo le analisi su incidenti reali documentati nel 2026.

Il percorso operativo consigliato dai principali practitioner segue tre fasi progressive. Si parte dall'OWASP Top 10 for Agentic Applications per costruire la comprensione di base delle categorie di rischio. Si progredisce verso il red teaming AI-specific, testando empiricamente le ipotesi del threat model attraverso attacchi controllati (prompt injection, context poisoning, tool misuse). Si struttura infine lo sviluppo sicuro con controlli specifici: validazione dell'input nei tool, principio del minimo privilegio per le autorizzazioni degli agenti, logging granulare delle azioni autonome, e meccanismi di human-in-the-loop per le azioni ad alto impatto.

7.3 Supply chain del software

Il Cyber Resilience Act europeo rende il threat modeling della supply chain un obbligo per i produttori di hardware e software con elementi digitali. Il modello di minaccia deve coprire l'intero ciclo di vita del componente: dalle dipendenze open source (con generazione e manutenzione del Software Bill of Materials) all'integrità degli artefatti di build, fino alla distribuzione sicura degli aggiornamenti. Nel

2026 sono stati identificati quarantatré componenti differenti di framework per agenti AI con vulnerabilità incorporate, a dimostrazione che la supply chain AI merita un sotto-dominio specifico del threat modeling.

7.4 OT, ICS e infrastrutture critiche

La convergenza IT-OT ha eliminato il tradizionale "air gap" che proteggeva i sistemi industriali. Il threat modeling per ambienti OT/ICS presenta sfide peculiari: i sistemi legacy non sono aggiornabili, i tempi di fermo per manutenzione sono limitati, e le conseguenze di una compromissione possono avere impatto fisico diretto sulla sicurezza delle persone. La convergenza tra crimine finanziario, minacce insider e attacchi cyber-fisici crea scenari in cui un attore può colpire simultaneamente la componente IT, la componente OT e i processi di business, con effetti a cascata che nessun threat model tradizionale è in grado di catturare integralmente.

Il riferimento metodologico per questo dominio è IEC 62443, standard internazionale per la sicurezza dei sistemi di automazione e controllo industriale, che deve essere integrato con MITRE ATT&CK for ICS per la mappatura delle tecniche di attacco documentate contro infrastrutture industriali.

7.5 Post-quantum readiness

Ogni sistema che gestisce dati con requisiti di confidenzialità a lungo termine è già oggi esposto al rischio "harvest-now-decrypt-later": gli attaccanti raccolgono traffico cifrato con la certezza di poterlo decifrare non appena i computer quantistici raggiungeranno capacità crittograficamente rilevanti. Il NIST ha completato nel 2024 la standardizzazione dei primi algoritmi post-quantum (CRYSTALS-Kyber per la cifratura a chiave pubblica, CRYSTALS-Dilithium per le firme digitali). Il threat modeling del 2026 deve includere la catalogazione degli algoritmi crittografici in uso nell'organizzazione — cosiddetto "crypto inventory" — e la prioritizzazione della migrazione verso PQC in base alla

sensibilità dei dati e alla loro durata di vita rilevante.

8. La prospettiva italiana ed europea: gap e opportunità

L'Italia si trova in una posizione di transizione nel 2026. Il recepimento di NIS2 (D.lgs. 138/2024) ha formalmente esteso l'obbligo di gestione del rischio cyber a migliaia di soggetti essenziali e importanti, molti dei quali non disponevano in precedenza di un processo formalizzato di threat modeling. L'Agenzia per la Cybersicurezza Nazionale ha pubblicato linee guida operative che richiamano esplicitamente l'adozione di metodologie strutturate di valutazione del rischio.

Il gap più critico riguarda le medie imprese del manifatturiero avanzato e delle filiere industriali: organizzazioni con ambienti OT significativi, spesso prive di un team di sicurezza dedicato, che devono costruire un processo di threat modeling praticamente da zero in tempi ristretti per la conformità NIS2. In questo contesto la scelta del framework e degli strumenti è determinante: un approccio troppo complesso rischia di essere abbandonato prima di produrre valore reale.

Le grandi organizzazioni italiane – banche, assicurazioni, utility, telco – sono generalmente più avanzate, con programmi di threat modeling maturi e in alcuni casi integrati nei cicli di sviluppo agile. Il gap in questo segmento riguarda prevalentemente la capacità di estendere il threat modeling ai sistemi AI che stanno deployando rapidamente, e la gestione della supply chain di fornitori e partner.

A livello europeo, il 2026 è l'anno in cui l'AI Act entra in piena applicazione per la maggior parte degli obblighi che riguardano i sistemi AI ad alto rischio. Le organizzazioni che implementano agenti AI autonomi devono dimostrare conformità con i requisiti di gestione del rischio, trasparenza e

supervisione umana — requisiti che si soddisfano attraverso un threat model documentato, non attraverso dichiarazioni generiche di conformità.

9. Gli errori più comuni da evitare

Il threat modeling mal praticato è peggio dell'assenza di threat modeling: crea una falsa percezione di sicurezza e consuma risorse senza produrre valore reale. Questi sono i pattern anti-threat-modeling più frequenti osservati nelle organizzazioni nel 2026.

Fare il threat model una volta sola. Il threat model è un artefatto vivente, non un documento da archiviare dopo la fase di design. Ogni cambiamento architetturale significativo, ogni nuova integrazione, ogni modifica del threat landscape esterno richiede una revisione. Le organizzazioni mature lo collegano direttamente al processo di change management: nessuna modifica architetturale rilevante viene approvata senza una valutazione dell'impatto sul threat model esistente.

Limitarsi alle vulnerabilità tecniche. Un threat model che si concentra esclusivamente su CVE, misconfiguration e debolezze del codice è sistematicamente cieco rispetto ai vettori più utilizzati nel 2026: social engineering dell'identità, abuso di accessi legittimi, supply chain compromise, manipolazione del contesto nei sistemi AI. Il modello deve rappresentare l'attaccante reale, non solo il sistema difeso.

Escludere gli stakeholder non tecnici. Il threat modeling efficace richiede contributi da team di sicurezza, engineering, product management, UX e legal/compliance. Escludere le funzioni non tecniche produce threat model tecnicamente accurati ma disconnessi dagli obiettivi di business e dai requisiti normativi. Un CISO legale presente nella sessione di threat modeling può identificare in pochi minuti obblighi di notifica o vincoli di data residency che

cambiano radicalmente la prioritizzazione delle minacce.

Non collegare il threat model ai controlli di sicurezza. Un threat model che produce una lista di minacce senza mappatura sui controlli esistenti e sulle lacune di copertura è un esercizio intellettuale, non uno strumento operativo. Il valore reale emerge quando il threat model diventa il driver delle decisioni di investimento in sicurezza: “investiamo in questo controllo perché copre queste tre minacce ad alta priorità del nostro threat model”.

Usare un solo framework per tutti i contesti. STRIDE è eccellente per applicazioni web tradizionali e pipeline DevSecOps; è inadeguato per sistemi AI agentici o ambienti OT industriali. L’errore opposto – adottare un framework complesso come PASTA per ogni contesto, inclusi i sistemi più semplici – produce overhead sproporzionato che porta all’abbandono del processo. La scelta del framework deve essere contestuale.

Ignorare i sistemi AI come componenti del threat model. Molte organizzazioni nel 2026 continuano a modellare le minacce ai propri sistemi senza includere i servizi AI integrati (LLM API, agenti, sistemi di raccomandazione). Questi componenti hanno superfici d’attacco specifiche – prompt injection, data poisoning, model extraction – che non vengono catturate dai framework classici e che possono diventare punti di ingresso per compromissioni più ampie.

10. Un percorso di maturità: da zero alla piena integrazione

Livello 0: punto di partenza (organizzazioni senza processo formalizzato)

Il primo obiettivo non è la perfezione metodologica ma l’abitudine. Si inizia con sessioni di threat modeling strutturate per i nuovi progetti o per le modifiche

architetturali significative, usando STRIDE come framework di riferimento per la sua accessibilità. Bastano tre ore con le persone giuste — un architetto, un developer, un responsabile della sicurezza — e un foglio bianco con il DFD dell'applicazione per produrre un primo threat model utile. L'output minimo è una lista di minacce prioritizzate e una prima mappatura sui controlli esistenti e mancanti.

Livello 1: standardizzazione (sei-dodici mesi)

Il processo viene formalizzato e reso ripetibile: template standard per i DFD, checklist metodologiche, criteri di prioritizzazione del rischio condivisi. Il threat modeling viene integrato nel processo di approvazione delle modifiche architetturali: nessuna modifica rilevante viene implementata senza una sessione di revisione del threat model. Si adottano i primi strumenti di supporto (OWASP Threat Dragon, Microsoft Threat Modeling Tool) per ridurre il tempo necessario per la documentazione.

Livello 2: integrazione DevSecOps (dodici-ventiquattro mesi)

Il threat modeling viene integrato nel ciclo di sviluppo agile: ogni epic o sprint significativo include una micro-sessione di threat modeling (trenta-sessanta minuti) focalizzata sulle modifiche specifiche. Si adottano piattaforme commerciali come IriusRisk o ThreatModeler per gestire l'inventario dei threat model, tracciare le remediation e generare reportistica per il management. Il threat model viene collegato al sistema di ticketing (Jira, Azure DevOps) in modo che ogni minaccia identificata generi automaticamente un work item di sicurezza tracciabile.

Livello 3: automazione e continuous threat modeling (oltre i ventiquattro mesi)

Il threat model viene generato e aggiornato semi-automaticamente a partire dai diagrammi architetturali

mantenuti nel repository di codice, con tool AI-assisted che suggeriscono minacce rilevanti sulla base delle modifiche apportate. Il threat modeling AI-specific (MITRE ATLAS, MAESTRO, OWASP Agentic Top 10) viene integrato per tutti i sistemi che includono componenti di intelligenza artificiale. Il threat model diventa la base documentale per la conformità regolamentare (NIS2, AI Act, CRA, DORA), eliminando la duplicazione tra processi di sicurezza e processi di compliance.

11. Il contesto regolamentare: la compliance come driver di threat modeling

L'ambiente regolamentare europeo nel 2026 è il più esigente della storia per le organizzazioni che operano nel digitale. Quattro framework normativi convergono simultaneamente in fase di enforcement.

La Direttiva NIS2 (recepita in Italia con D.lgs. 138/2024) impone la formalizzazione di un processo di gestione del rischio che include esplicitamente l'analisi delle minacce alle reti e ai sistemi informativi, con audit regolari e notifica delle vulnerabilità significative entro ventiquattro ore per gli incidenti con impatto significativo. Il threat modeling strutturato è lo strumento naturale per soddisfare questo requisito in modo documentabile e verificabile.

L'AI Act dell'Unione Europea ha portato in piena applicazione, nell'agosto 2026, la maggior parte degli obblighi per i sistemi AI ad alto rischio. Le organizzazioni che deploiano agenti AI autonomi devono dimostrare conformità con i requisiti di gestione del rischio, trasparenza e supervisione umana. Un threat model documentato dei sistemi AI – che copra le minacce specifiche all'autonomia, alla memoria e alla catena degli strumenti – è la base tecnica per questa dimostrazione.

Il Cyber Resilience Act rende obbligatorio il Software Bill of Materials e impone requisiti di sicurezza lungo tutto il ciclo di vita del prodotto digitale, inclusa la gestione delle vulnerabilità e la disponibilità di aggiornamenti sicuri. Il threat model della supply chain software diventa parte integrante del dossier tecnico richiesto per la conformità.

DORA (Digital Operational Resilience Act) estende analoga disciplina al settore finanziario, con particolare enfasi sui Threat-Led Penetration Test: test di penetrazione avanzati basati su threat intelligence reale, condotti da fornitori certificati, che richiedono come prerequisito un threat model maturo delle infrastrutture critiche dell'organizzazione.

Le figure di Legal e Compliance devono essere parte integrante del processo di threat modeling, non destinatarie passive dei suoi output. La loro presenza nelle sessioni di analisi garantisce che vengano identificate e modellate minacce con implicazioni normative specifiche: violazioni di data residency, scenari di data breach soggetti a notifica obbligatoria, rischi reputazionali connessi a incidenti che coinvolgono categorie particolari di dati.

12. Raccomandazioni operative per il 2026

Adottare un approccio multi-framework calibrato sul contesto. Nessun singolo framework copre tutti i domini rilevanti nel 2026. La combinazione più solida è: STRIDE o PASTA come metodologia base per i sistemi software tradizionali, MITRE ATLAS più OWASP Agentic Top 10 per i sistemi AI, LINDDUN per i componenti con trattamento di dati personali, IEC 62443 per gli ambienti OT industriali. La scelta non deve essere ideologica ma pragmatica: il framework migliore è quello che il team effettivamente usa e aggiorna.

Trattare l'identità come superficie primaria. Il threat model deve riflettere il nuovo paradigma degli attaccanti: la

compromissione inizia quasi sempre attraverso l'identità, non attraverso le vulnerabilità tecniche. Questo significa modellare esplicitamente gli scenari di abuso di accessi legittimi, session hijacking post-MFA, exploitation delle integrazioni SaaS e social engineering degli helpdesk IT. La mitigazione più efficace – continuous identity verification e routing di tutte le applicazioni SaaS attraverso un identity provider centralizzato – deve essere guidata dal threat model, non dall'intuizione.

Integrare il threat modeling nel ciclo di vita degli agenti AI dall'inizio. I sistemi agentici richiedono modelli di sicurezza che comprendano obiettivi, strumenti, contesto e flusso decisionale. I controlli tradizionali di rete, IAM e sicurezza applicativa rimangono necessari ma non più sufficienti. Il logging granulare delle azioni autonome, il principio del minimo privilegio per le autorizzazioni degli agenti e i meccanismi di human-in-the-loop per le azioni ad alto impatto devono essere definiti nel threat model prima del deployment.

Mappare il threat model sui requisiti normativi. AI Act, NIS2, CRA e DORA richiedono documentazione formale del processo di gestione del rischio. Un threat model strutturato correttamente diventa la base documentale per la conformità regolamentare, eliminando la duplicazione tra processi di sicurezza e processi di compliance. Questa sinergia riduce significativamente il costo complessivo della conformità.

Passare dagli indicatori di compromissione statici al rilevamento comportamentale. Con gli attaccanti che cambiano rapidamente l'infrastruttura e distribuiscono malware custom eseguito in memoria, affidarsi esclusivamente agli IoC statici non è più sufficiente. Il threat model deve anticipare questi scenari comportamentali e guidare l'implementazione di modelli di rilevamento basati sulle anomalie rispetto alle baseline stabilite.

Avviare il crypto inventory e la roadmap post-quantum. Ogni sistema che gestisce dati con requisiti di confidenzialità a lungo termine merita una sezione specifica del threat model dedicata ai rischi crittografici. Il primo passo pratico è il censimento degli algoritmi crittografici in uso nell'organizzazione; il secondo è la prioritizzazione della migrazione verso algoritmi post-quantum in base alla sensibilità dei dati e alla loro durata di vita rilevante.

Costruire un team interdisciplinare e mantenerlo nel tempo. Il threat modeling efficace non è un'attività che si delega a un singolo esperto di sicurezza. Richiede la partecipazione attiva di architetti software, sviluppatori, product manager, responsabili legali e compliance officer. Il valore del processo cresce con la maturità del team: un'organizzazione che pratica il threat modeling da due anni produce output qualitativamente superiori rispetto a una che lo ha appena avviato, a parità di metodologia e strumenti.

Conclusione

Il threat modeling nel 2026 non è più un esercizio teorico confinato alla fase di design del software. È una disciplina operativa continua, regolamentariamente necessaria e tecnicamente inedita nella sua complessità. La combinazione di attacchi che si industrializzano, sistemi AI autonomi che moltiplicano la superficie d'attacco e un quadro normativo europeo in piena enforcement rende questa disciplina non rinviabile per nessuna organizzazione che gestisca sistemi digitali critici.

Il paradosso del threat modeling è che le organizzazioni che ne hanno più bisogno – quelle con meno maturità in sicurezza, spesso PMI o soggetti appena rientrati nel perimetro NIS2 – sono anche quelle più tentate di rimandarlo, percependolo come troppo complesso o troppo costoso. La risposta è che il costo di un processo di threat modeling ben calibrato è sistematicamente inferiore al costo medio di una singola

violazione: non come stima teorica, ma come dato empirico confermato da anni di ricerca sul ROI della sicurezza preventiva.

Le organizzazioni che hanno già investito in processi maturi di threat modeling si trovano nel 2026 in una posizione di vantaggio strutturale: dispongono di una difesa più efficace, di una documentazione di conformità che i regolatori richiedono esplicitamente, e di una capacità di risposta agli incidenti più rapida perché i percorsi di attacco più probabili sono già stati anticipati. Per chi non ha ancora strutturato questo processo, il momento per iniziare – con umiltà metodologica, strumenti proporzionati alla propria maturità e un team interdisciplinare – è adesso.

Condividi sui Social Network:
https://www.ictsecuritymagazine.com/notizie/threat-modeling-2026/