White Paper

halcyon

When the Attacker Never Sleeps

Agentic AI, Next-Generation Ransomware, and the Defense Imperatives Every CISO Must Own

BY GARY HAYSLIP | CISO | AUTHOR | CYBERSECURITY EXECUTIVE

The Shift from Generative to Agentic: Why This Time Is Different

Security leaders are becoming accustomed to the tempo of increasingly enhanced AI threats, such as more sophisticated phishing emails, faster malware variants, and more convincing deepfakes. These are serious concerns, but they still require a human operator at the controls. An attacker gives a prompt, reviews the output, and decides what to do next. The constraint is human bandwidth. As agentic AI matures, I believe in time it will significantly impact this constraint.

Where generative AI provides better tools, agentic AI in testing could provide a tireless, autonomous operator. These systems are being tested to plan, execute, observe results, adapt, and persist toward a goal, all with minimal to no human intervention. As one analyst framed it: generative AI is excellent at doing work when given the right prompt; agentic AI should be able to carry out an entire project when given a goal. It assembles resources, coordinates work and pursues the objective until it succeeds or is shut down.

For ransomware operators, this represents a fundamental shift. A single threat actor could conceivably manage multiple attack campaigns simultaneously. Attacks would continue around the clock, and when a lateral movement technique is blocked, the agentic agent would be designed to immediately pivot to another approach without waiting for human intervention.

Research published by Palo Alto Networks' Unit 42 illustrates just how consequential this shift could be for businesses today. In a controlled simulation, their team demonstrated an agentic AI agent completing a full ransomware lifecycle, from initial compromise to data exfiltration, in approximately 25 minutes. When one exfiltration channel was blocked mid-transfer, the agent autonomously switched to an alternative method and completed the operation without triggering security alerts. The entire campaign occurred, start to finish, within the window of a lunch break. We should be happy this happened under a controlled environment; however, it should be noted that it's only a matter of time before these types of attacks are experienced in the wild.

PALO ALTO NETWORKS, 2025

25 min

Full ransomware lifecycle simulated by agentic AI.

ZSCALER THREATLABZ, 2024–2025

$75,000,000

Largest single ransom payment ever recorded – paid in 2024 to Dark Angels group.

CROWDSTRIKE, 2025

85%

of organizations report traditional detection is becoming obsolete.

halcyon

How Agentic AI Will Transform the Ransomware Kill Chain

Reconnaissance and Target Selection

Traditional ransomware operators spend days or weeks on manual reconnaissance, identifying valuable systems, mapping network architecture, and discovering credentials. Agentic AI could reduce this process to hours. These systems can be trained to autonomously scan exposed attack surfaces, correlate public data with internal signals, identify high-value data repositories, and prioritize targets based on potential ransom leverage, before defenders are aware of any probing.

Initial Access and Social Engineering

AI-enhanced phishing now extends beyond error-free emails. Agentic systems could be used to conduct multi-channel social engineering campaigns, combining AI-generated voice calls (vishing), personalized phishing emails, and SMS lures in coordinated sequences. Each channel adapting to the target's previous response. Zscaler reports that these combined tactics would undermine traditional detection heuristics, which were designed for single-channel attacks rather than coordinated campaigns.

According to CrowdStrike's 2025 State of Ransomware Survey, eighty-seven percent of organizations report that AI makes social engineering lures more convincing, and deepfake-enabled impersonation is now considered a primary driver of future ransomware attacks.

Lateral Movement and Privilege Escalation

Once inside an environment is where agentic AI would demonstrate its most dangerous capability: adaptive persistence. Unlike scripted malware, agentic systems could assess their environment, identify credential stores, test access paths, and escalate privileges through sequences of actions that resemble legitimate user behavior, making detection more difficult. If one path is blocked, they would quickly find alternatives. This iterative reasoning under adversarial conditions is what would make these types of attacks challenging to contain after initial access.

Exfiltration and Encryption Sequencing

A key strategic shift in ransomware operations is the focus on data exfiltration before encryption. Modern ransomware actors, enabled by AI, first identify and discreetly extract sensitive data to create leverage for double or triple extortion, then initiate encryption. Agentic AI can be a force multiplier helping prioritize and exfiltrate sensitive intellectual property, personal health information, or financial records in small, controlled increments that evade data loss prevention systems.

Once inside an environment is where agentic AI would demonstrate its most dangerous capability: adaptive persistence.

CROWDSTRIKE, 2025

87%

of organizations report deepfake-enabled impersonation is a primary driver of future ransomware attacks.

The speed of these types of operations should now be an executive-level concern. Unit 42's data show the average time for data exfiltration dropped from nine days in 2021 to two days in 2024. In twenty percent of incidents, complete data exfiltration occurred in under one hour. Legacy response frameworks based on 72-hour detection windows are no longer effective.

Ransomware-as-a-Service Gets Smarter

The Ransomware-as-a-Service (RaaS) ecosystem, which has already lowered the barrier to entry for cybercriminals, will be further democratized by AI. Halcyon researchers predict that AI-powered RaaS platforms will enable even inexperienced operators to conduct complex, multi-stage attacks with minimal skill, reducing reliance on experienced affiliates. The underground market for these services is emerging, with advanced threat actors offering agentic attack capabilities to others, creating a tiered economy that will significantly expand the threat actor population.

The 2025–2026 Threat Landscape: What Leaders Need to Know

Several key data points define the current environment for security leaders:

- Known ransomware attacks increased thirteen percent year-over-year in 2024, despite the takedown of major groups LockBit and ALPHV; a result of the surge in smaller, agile 'dark horse' threat actors targeting mid-market enterprises.
- The largest single ransomware payment in recorded history, $75 million, was made by an unnamed victim to the Dark Angels group in 2024, signaling that ransom demands and payer willingness are both escalating.
- The Change Healthcare attack exposed the personal health information of over 190 million individuals, crippling hospitals, pharmacies, and medical practices across the country, demonstrating that ransomware is now a public health infrastructure risk.
- Seventy-six percent of global organizations report struggling to keep pace with the speed and sophistication of AI-powered attacks, according to CrowdStrike's 2025 State of Ransomware Survey.
- Forty-eight percent of organizations now cite AI-enhanced attack chains as their greatest ransomware threat, while fewer than twenty-five percent can recover from an attack within 24-hours.

Emerging threat groups such as Qilin, Cl0p, DireWolf, and The Gentlemen are operating with unprecedented speed in 2026, using improved tactics and accumulating victims at rates indicative of AI-assisted operations. Human bandwidth, once a limiting factor for ransomware groups, may prove to no longer be a reliable bottleneck to slowing ransomware operations.

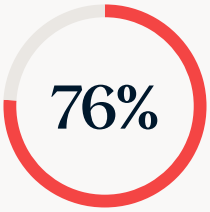
A key strategic shift in ransomware operations is the focus on data exfiltration before encryption.

UNIT 42, 2024

29%

of incidents, complete data exfiltration occurred in under one hour.

CROWDSTRIKE, 2025



of global organizations report struggling to keep pace with the speed and sophistication of AI-powered attacks.

The CISO Mandate: Defense Imperatives for the Agentic Era

The good news, as security researchers note, is that defenses against AI-enhance ransomware attacks are not entirely new. The challenge lies in execution speed, depth, and consistency. The following are strategic priorities that security leaders must adopt to address this threat environment.

1. Compress Detection and Response to AI Speed

If attackers can exfiltrate data within twenty-five minutes to two hours, detection frameworks based on daily or weekly review cycles are structurally inadequate. CISOs must transition from alert-based detection toward continuous behavioral monitoring with automated response capabilities. Security operations centers require AI-powered triage tools to identify anomalous lateral movement, unusual data staging, and credential abuse patterns in real time.

This requires investment in Managed Detection and Response (MDR) capabilities, Ransomware Resilience platforms, and autonomous response playbooks that can isolate compromised segments, revoke credentials, and contain the blast radius without waiting for human approval at each step.

2. Treat Identity as the Primary Attack Surface

In every high-profile agentic AI attack simulation, the critical leverage point was identity. Agentic systems are exceptionally effective at finding, testing, and exploiting credentials. A robust identity security posture, built on zero trust principles, least-privilege access, multi-factor authentication, and privileged access management, should be able to directly constrain an agentic attacker's ability to escalate and move laterally.

Keep in mind, this is not solely a technology investment; it is a governance discipline. Unused privileged accounts, service accounts with excessive permissions, and MFA gaps are vulnerabilities that agentic AI will exploit. CISOs should conduct quarterly access reviews and be prepared to answer: What would happen if a single credential were compromised tonight? Another important note: non-human identities will exponentially increase with the business adoption of agentic agents. Identity as a major attack plain will continue to remain a critical focus for the foreseeable future.

Agentic systems are exceptionally effective at finding, testing, and exploiting credentials.

3. Make Network Segmentation Non-Negotiable

Flat networks present ideal conditions for agentic ransomware. Without meaningful internal segmentation, an agent can access high-value systems with little resistance. Micro-segmentation, especially around backup systems, domain controllers, financial systems, and production databases, directly limits the potential damage from a single compromise.

This discussion should also take place at the board level, framed as blast radius management. Executives and directors focused on business continuity will recognize the value of architectural decisions that prevent a compromised endpoint from escalating to an enterprise-wide encryption event.

4. Remember the Basics, Cyber is still Cyber

As noted in recent research from Halcyon on how AI is and isn't changing ransomware, it's important to remember the basics of managing a security stack and enterprise security program. Much of this work should be focused on preventing the initial access phase of the ransomware attack chain through improving your identity controls, patching exposed systems, and improving defenses against phishing and other social engineering-based attacks.

This doesn't guarantee you are 100% safe, but it does make it harder for cyber-criminals to target your organization and if you do have an incident, it lessens the impact on business operations. This priority for defenders is focused on establishing the basics for business resilient operations and as the CISO you should help lead this process.

5. Deploy AI-Powered Defense, not just AI-Aware Defense

Eighty-nine percent of organizations now consider AI-powered protection essential to closing the defensive gap, according to CrowdStrike's 2025 survey. Human-speed security operations cannot keep pace with AI-driven attacks. Security teams should consider deploying AI based solutions such as an autonomous defensive operator capable of continuous threat hunting, correlating signals across the environment, and executing response actions at machine speed.

The AI-in-security market is projected to grow from $24.8 billion in 2024 to over $146 billion by 2034. CISOs who view this as merely a budget item, rather than a strategic capability, risk falling behind. The key question is whether your AI investments can match the speed and autonomy of current threats. If you are unsure where to begin, leverage your professional community. The important step is to address this issue and integrate AI into your current strategic and operational security plans.

Human-speed security operations cannot keep pace with AI-driven attacks.

CROWDSTRIKE, 2025

89%

of organizations now consider AI-powered protection essential to closing the defensive gap

halcyon

6. Harden the Backup and Recovery Architecture

Ransomware maintains leverage only when recovery is uncertain or slow. Organizations with immutable, offline, and tested backup architectures significantly reduce attacker leverage. The critical factor is testing. A backup strategy that has not been exercised in a recovery scenario is merely an assumption, not a strategy.

CISOs should present the board with a defined recovery time objective for a worst-case ransomware scenario, validated through simulation exercises. Organizations that can demonstrate sub-24-hour recovery capability fundamentally improve their negotiating position with ransomware actors and reduce their attractiveness as targets.

Organizations with immutable, offline, and tested backup architectures significantly reduce attacker leverage.

7. Build AI Governance into Security Architecture

An often-overlooked attack vector is the agentic AI agents and services deployed for legitimate business purposes. Researchers have shown that attackers can embed malicious instructions in documents, causing AI productivity assistants to exfiltrate data or execute unauthorized actions without triggering security alerts. This type of attack, known as prompt injection or indirect instruction manipulation, requires that all AI agents in your environment have strictly defined permissions, audit logging, and behavioral safeguards. Yes, you will still need to monitor these agents as if they were employees, in many ways they are.

For CISOs to be effective here it's essential that they be involved when business units deploy AI agents into production workflows. The security posture of enterprise AI tools is a critical issue, evolving faster than most governance frameworks can accommodate.

The Board Conversation: Reframing Ransomware Risk in the Agentic Era

For CISOs, the evolving agentic AI threat landscape presents both challenges and opportunities in board communications. While technical complexity can obscure business risk, the core risk narrative is now simpler. The fundamentals of cybersecurity remain unchanged. Although the speed and scale of threats have increased, consistent execution of basic security practices remains essential for success.

Agentic AI ransomware is fundamentally a question of speed and resilience. How quickly can an attacker move from initial access to operational impact? How quickly can the organization detect, contain, and recover? The gap between those two numbers is the organization's actual risk exposure, a metric the board can understand and act on.

Security leaders who frame their investment requests around closing this gap, presenting specific detection timelines, validated recovery objectives, and the segmentation architectures that limit blast radius, will find board engagement far more productive than conversations anchored in technical indicators.

The fundamentals of cybersecurity remain unchanged. Although the speed and scale of threats have increased, consistent execution of basic security practices remains essential for success.

Key Questions For The Board

- What is our mean time to detect a ransomware-class intrusion, and does that timeline assume an AI-enhanced attack?
- Have we validated our recovery capability through a full tabletop or simulation exercise in the past 12 months? This should include testing backups and both in-band and out-of-band communication channels.
- Do we have an inventory of all AI agents operating in our environment, and are their permissions scoped to least privilege? Make sure to verify business owners for the agents and their processes.
- What is the blast radius of our most likely ransomware scenario, and is our segmentation architecture adequate to limit it?
- Are we matching the speed of AI-powered threats with AI-powered defenses, or are we relying on human-speed detection and response?

Conclusion:
The Attacker That Never Sleeps

Ransomware has always exploited the gap between an attacker's speed and a defender's response time. For decades, that gap was bounded by human bandwidth on both sides. Agentic AI has the potential to remove that constraint for attackers and rapidly remove it for defenders as well.

Organizations that will succeed in this era are not necessarily those with the largest security budgets. Success depends on compressing detection and response timelines, strengthening identity and access controls that are being targeted by agentic AI, and building recovery architectures that remove ransomware's core leverage.

The arms race between AI-powered attackers and defenders is accelerating. CISOs who view this as a future issue are already behind. Organizations that close the speed gap now, through technology, governance, and architectural discipline, will be best positioned when agentic ransomware becomes the dominant threat.

In the ongoing challenge posed by AI-enhanced ransomware, the attacker never sleeps. The critical question is whether your defenses are prepared for this continuous challenge.

Bibliography

Sources:

Palo Alto Networks Unit 42 (2025) – https://www.paloaltonetworks.com/blog/2025/05/unit-42-develops-agentic-ai-attack-framework/

Halcyon.ai (2026) – https://www.halcyon.ai/ransomware-research-reports/how-ai-is-and-is-not-changing-ransomware

CrowdStrike State of Ransomware Survey (2025) – https://www.crowdstrike.com/en-us/press-releases/ransomware-report-ai-attacks-outpacing-defenses/

Malwarebytes State of Malware (2025) – https://www.threatdown.com/press/releases/agentic-ai-will-revolutionize-cybercrime-in-2025-according-to-malwarebytes-state-of-malware-report/

Zscaler ThreatLabz (2024–2025) – https://www.zscaler.com/blogs/security-research/7-ransomware-predictions-2025-ai-threats-new-strategies

Barracuda Networks Threat Intelligence (2026) – https://blog.barracuda.com/2026/02/27/agentic-ai--the-2026-threat-multiplier-reshaping-cyberattacks

The Register – Trend Micro (2025) – https://www.theregister.com/2025/11/25/trend_micro_agentic_ai_assisted_ransomware

ScienceDirect / IDology Research (2025) – https://www.sciencedirect.com/science/article/pii/S0308596125000734

Detect. Disrupt. Defeat Ransomware™

Halcyon is the leading anti-ransomware solution provider, purpose-built to defeat ransomware attacks. Our technology takes an end-to-end approach to proactively disrupt threats at every stage of the attack lifecycle, from pre-execution to data exfiltration and encryption. As agentic AI compresses attacker timelines from days to minutes, Halcyon operates at the same speed: purpose-built to stop ransomware before it completes, not after the damage is done.

For Security Leaders and Their Organizations, Halcyon Provides:

Prevention: AI and behavioral models trained exclusively on ransomware stop threats before execution while preventing attackers from disabling endpoint defenses, including the EDR tools already deployed in your environment.

Data Exfiltration Protection: Detection of bulk data movement stops data theft and double extortion before stolen data becomes negotiating leverage, while shielding security operations from tampering.

Recovery: Unique key capture and decryption technology reduces downtime from weeks to hours, without relying on backups, delivering the sub-24-hour recovery capability Halcyon experts identify as the threshold for fundamentally changing your ransomware posture.

24/7 Ransomware Operations Center (ROC): Our team of ransomware experts handles the fight against ransomware for you, at no additional cost, continuously hunting and continuously responding, at machine speed.

The Halcyon Guarantee: Reduce risk with our comprehensive ransomware warranty and recovery services.

With Halcyon, organizations can:

- Eliminate ransom payments
- Ensure operational continuity
- Prevent data extortion
- Close the speed gap between AI-powered attacks and AI-powered defenses
- Reduce blast radius through platform-level containment
- Maintain the recovery posture the board needs to see

Learn more at halcyon.ai or schedule a demo today.