knowbe4

The New Face of Ransomware: How AI and Exfiltration Are Mutating the Threat Landscape

WWW.KNOWBE4.COM

KnowBe4, Inc. | 33 N Garden Ave, Suite 1200, Clearwater, FL 33755
Tel: 855-KNOWBE4 (566-9234) | Email: Sales@KnowBe4.com

CR08590626

Introduction

Over the past year, something has shifted in how security leaders are talking about ransomware. The topic is coming up less often in boardroom briefings. It's also generating fewer headlines. Sometimes it seems as though it's dropped from the collective radar.

If so, that's a mistake. The ransomware threat is still very much with us, but it has changed shape. The force transforming it is artificial intelligence (AI), which has made attacks faster and cheaper. Even more alarming, AI has removed the skill floor for attackers. Campaigns that once demanded genuine technical expertise now require only a subscription.

That shift has implications for organizations of all sizes and sectors. This article explores how the ransomware landscape has changed, how AI is mutating the threat in ways that matter for your defenses, why your people remain central to any credible response, and what a modern human-AI defense looks like in practice.

The Changing Ransomware Landscape

The classic encryption model most people still think of when they hear the word “ransomware” is fading. The extortion model is thriving in its place, and the numbers tell the story.

According to Coveware, data exfiltration — the theft of sensitive files from a victim’s network — now appears in 76% of ransomware cases. In many incidents, data exfiltration is the primary objective rather than a secondary pressure tactic. At the same time, ransom payments are dropping. The average ransom payment fell to USD $376,941 in Q3 2025, down 66% from the prior quarter, and the median dropped to USD $140,000, a 65% decline over the same period.

At first glance, such declines may appear to be wins for defenders. They’re not. Rather, they reflect adaptation. Although victims are paying less, they remain vulnerable in ways that matter more than ever.

A growing share of Coveware’s Q3 2025 cases involved exfiltration with no encryption at all. In that category, ransom payment rates fell to a record low of 19%, suggesting victims no longer believe paying will stop their data from being released. It seems that more organizations are choosing to absorb the breach rather than fund the attacker. That’s a choice with its own downsides in the form of published data, regulatory exposure, and lasting reputational damage.

For organizations in regulated industries, the consequences compound. A publicized breach of protected health information or personally identifiable data triggers reporting obligations and legal exposure on top of the reputational damage — costs that can dwarf the original downtime.

Strong backups — once considered the ultimate insurance policy against ransomware — are no longer enough on their own. The threat has shifted from “we locked your systems” to “we have your data.” No restore process can fix the latter.

A Fracturing, Reforming Ecosystem

Amid changing ransomware tactics, the ecosystem itself has been churning. To their credit, law enforcement has disrupted several major gangs over the past few years. But survivors remain, and some have consolidated into newly formed super-syndicates and alliances. The result is an unexpected combination of greater distribution among smaller actors alongside greater coordination among the largest players.

At the same time, we're witnessing the rise of an access broker economy, in which corporate network access is sold to the highest bidder via dark web auctions. Unfortunately, the price of entry has dropped dramatically. According to Darkweb IQ as cited by Chainalysis, the average price for victim access fell from about USD $1,427 in Q1 2023 to just USD $439 in Q1 2026.

In plain terms, initial access to a corporate network now costs about as much as a budget smartphone. That price collapse is a direct consequence of automation, and it pulls a new class of low-skill actors into the ecosystem.

AI is Mutating the Threat

As dropping access-broker prices suggest, AI has upended the economics of cybercrime broadly, and ransomware operators are borrowing freely from the playbook.

Chainalysis has reported that AI-enabled crypto scams are now 4.5x more profitable than traditional scams, extracting an average of USD $3.2 million per operation. The same underlying capabilities — including AI-generated lures, phishing-as-a-service, and automated personalization — are being ported directly into ransomware operations.

The most immediate change has been to phishing, which remains the entry point for most ransomware attacks. As noted in Microsoft's 2025 Digital Defense Report, AI automation has the potential to increase phishing profitability by as much as 50 times by scaling highly targeted attacks to thousands of victims at minimal cost. According to the same report, AI-generated lures produced click-through rates of 54%, compared to 12% for traditional emails. That represents a fundamental shift in attack economics.

AI-enabled attacks are especially hard to defend against because of how quickly they move. Mandiant's M-Trends 2026 report documents that attackers are speeding up their internal hand-offs dramatically. Across ransomware incidents broadly, payload deployment now routinely happens within hours of initial access rather than days or weeks later.

Consider what that means for your incident response (IR) plan. Most IR playbooks were built around an assumption of having time to detect, triage, and contain. When you have hours instead of days, most of those playbooks are structurally mismatched to the threat.

The trajectory points somewhere even more unsettling: agentic AI. Threat actors are beginning to deploy self-directed systems that do more than automate individual attack steps. These systems plan and execute entire campaigns end to end. They can adjust to network defenses, swap payloads mid-attack, and learn from detection responses in real time.

The evidence clearly shows that most hacking will be AI-enabled by the end of 2026. Several capabilities are already in use: AI-powered polymorphic malware that constantly mutates its signature to evade endpoint detection, LLM-enabled malware that has moved from proof-of-concept to practice, and deepfake technology being used to impersonate executives in social engineering attacks, coerce payments, or manufacture reputational pressure.

These capabilities are no longer theoretical threats discussed at conferences. They are showing up in incident reports.

Building a Human-AI Defense

AI-powered automatic personalization of phishing is now common. The traditional signals employees were taught to watch for, such as poor grammar, generic salutations, and implausible urgency, have been systematically eliminated from modern attacks.

Yet even as humans remain vulnerable, they also serve as an essential detection layer that no technical control can fully replace. A trained and skeptical employee catches what the tools miss. The goal is to make people the last and most lasting line of defense.

These five pillars form the foundation of an effective human-AI defense:

1. Multi-factor and Non-Phishable Authentication

Use MFA wherever possible and favor non-phishable options like hardware or app-based tokens. Avoid SMS-based MFA where possible and train users to always verify authentication requests they did not initiate. This single step alone can block a significant share of credential-theft attacks.

2. Email Authentication Standards

Implement DMARC (Domain-based Message Authentication, Reporting, and Conformance), DKIM (DomainKeys Identified Mail), and SPF (Sender Policy Framework) to validate sender identities and prevent spoofing attempts before they reach user inboxes. Although these protocols have been available for years, too many organizations have yet to fully deploy them.

3. AI vs. AI

Deploy artificial intelligence defensively within your email gateways and user-behavior analytics tools. AI can detect anomalies and suspicious patterns faster than humans or static rule sets. In an environment where attackers have automated everything, defenders who have not done the same are bringing a rulebook to a machine-speed fight.

4. Security Awareness and Ongoing Training

Equip employees with the knowledge to spot polymorphic phishing attempts. Continuous, adaptive training and simulated phishing exercises help keep users alert and resilient. Annual check-the-box training will not hold up against the volume and sophistication of today's attacks.

5. Zero Trust Culture

Educate users to assume every message, link, and request could be malicious. Healthy skepticism, backed by strong authentication and AI support, creates a human-AI partnership that can keep pace with modern threats.

None of these pillars is new; what has changed is the urgency of integrating them. Any one of them in isolation is inadequate against the current threat landscape. Together, they form a layered defense where the weaknesses of one layer are covered by the strengths of another.

The War Is Quieter, Not Over

Ransomware may have lost prominence in the headlines, but that's not because the war has been won. It has fallen off the radar because it has mutated into something harder to see. The activity and damage continue. The mechanism has changed.

Ransomware is becoming increasingly autonomous as AI automates the work behind it, including exploitation, reconnaissance, data analysis, and even negotiations. The organizations that weather this new landscape will be the ones that focus on resilience as well as prevention. Resilience, in this new reality, requires human and technical capabilities working in concert.

A reasonable starting point is to assess where your human layer stands today. That is where most ransomware attacks still begin, and it's where you can tap into your best chance to stop them.

Learn More about how KnowBe4 can secure the human/AI workforce against modern ransomware

Free Phishing Security Test
Find out what percentage of your employees are Phish-prone with your free Phishing Security Test

Free Email Exposure Check
Find out which of your users emails are exposed before bad actors do

Free Automated Security Awareness Program
Create a customized Security Awareness Program for your organization

Free Domain Spoof Test
Find out if hackers can spoof an email address of your own domain

Free Phish Alert Button
Your employees now have a safe way to report phishing attacks with one click

About KnowBe4

KnowBe4 empowers the modern workforce to make smarter security decisions every day. Trusted by more than 70,000 organizations worldwide, KnowBe4 is the pioneer of digital workforce security, securing both AI agents and humans. The KnowBe4 Platform provides attack simulation and training, collaboration security, and agent security powered by AIDA (Artificial Intelligence Defense Agents) and a proprietary Risk Score. The platform leverages 15-years of behavioral data to combat advanced threats including social engineering, prompt injection, and shadow AI. By securing humans and agents, KnowBe4 leads the industry in workforce trust and defense.

More information at www.KnowBe4.com.

Follow us on Linkedin

Follow us on X

knowbe4

KnowBe4, Inc. | 33 N Garden Ave, Suite 1200, Clearwater, FL 33755
855-KNOWBE4 (566-9234) | www.KnowBe4.com | Sales@KnowBe4.com