

How AI Reshaped Cybersecurity in 2025 and What 2026 Will Demand

Cybercriminals and defenders both leveraged artificial intelligence in 2025, turning cybersecurity into an AI-powered “arms race.”

2025 will be remembered as a pivotal year when artificial intelligence (AI) rewrote the playbook for both cyber attacks and cyber defenses. After years of hype, generative AI became a practical tool that cybercriminals weaponized and security teams adopted in response. The result was an unprecedented AI arms race in cybersecurity, with both sides using advanced AI to outsmart the other.

This brief examines:

- 🕒 How 2025 changed the rules of engagement
- 🕒 How criminals used generative AI to launch more convincing and scalable attacks
- 🕒 How AI transformed threat detection and response for defenders
- 🕒 The key events that shaped this new landscape
- 🕒 What security leaders should do heading into 2026

AI-POWERED ATTACKS

CYBERCRIMINALS EMBRACE GENERATIVE TECH

Cybercriminals have always exploited emerging tech but in 2025, they fully weaponized generative AI to accelerate, scale, and automate their attacks. LLMs (Large Language Models), like OpenAI's GPT and others, allowed attackers to learn faster, craft attacks instantly, and expand fraud operations beyond previous limits. Tasks that once demanded skill, labor, and time (e.g., phishing emails or deepfake videos) could now be created in seconds using a prompt. AI fundamentally lowered the barrier to sophisticated cybercrime.

Here are the most impactful ways criminals leveraged AI in 2025:

Flawless Phishing at Unprecedented Scale

Generative AI turned phishing into a whole new ballgame. Instead of clumsy emails riddled with grammar mistakes, attackers now send out polished, perfectly worded messages tailored to each victim.

AI models can scrape a target's online presence and craft highly personalized phishing emails that reference real projects, colleagues, or interests, making the scam far more believable. The results have been alarming. Industry reports noted an explosion in phishing volume (one analysis cited a [1,265% surge](#) in AI-linked phishing attacks) and sky-high success rates, with some AI-generated phishing emails tricking over half of recipients into clicking. Comparatively, generative AI needed only five minutes to design a phishing campaign as effective as one that took human experts 16 hours.

With AI, criminals can efficiently pump out uniquely tailored phishing lures at scale by the thousands to overwhelm traditional email defenses.

Deepfake Fraud and Impersonation Scams

2025 saw deepfake technology employed by cybercrime. “Deepfakes,” AI-generated synthetic media, allowed criminals to realistically impersonate voices and faces. In one high-profile incident at the global firm, Arup, fraudsters cloned the video and voice of the company’s CFO and other colleagues to conduct a live video call with an employee. The deepfake was convincing enough to trick the employee into transferring \$25.6 million to the attackers. This was not a network “hack” at all; it was technology boosted social engineering designed to breach human trust. And this was not an isolated case.

[Security surveys indicated that over 60% of organizations](#) encountered a deepfake-based attack attempt in the past year resulting in \$350 million in losses in just one quarter of 2025. This demonstrates how generative AI provides a powerful new toolkit for impersonation, letting attackers literally put words in someone else’s mouth to win trust and loot money.

Automated Social Engineering and Chatbot Scams

Beyond emails and videos, generative AI supercharged call-center fraud and text scams.

AI chatbots were deployed by fraudulent call centers to automate the initial stages of phone scams. These malicious bots mimicked human customer service, complete with friendly voices and real-time conversational responses guided by AI models.

The bots would engage victims, identify gullible targets, and then hand them off to human scammers to complete the con. This significantly increased the efficiency of fraudulent call operations.

Smishing (SMS phishing) also got an AI upgrade. Instead of generic “Your package is delayed, click here” texts, criminals blasted out SMS messages written by AI in flawless language, sometimes dynamically adjusting messaging to boost click-through rates. In short, AI became the scammer’s tireless assistant, churning out convincing lies across email, voice, and text channels.

Malware Generation and AI-Boosted Hacking

Generative AI didn’t just help with social tricks, it also assisted on the technical side of cybercrime. 2025 saw the rise of AI-generated malware and tools that can adapt in real-time. For example, attackers leveraged AI to create polymorphic malware — malicious code that constantly rewrites itself to evade detection. Some advanced strains were able to morph every 15 seconds during an attack, producing endless variations that signature-based antiviruses couldn’t identify. An estimated 70%+ of major breaches involved some form of polymorphic malware, [showing](#) how quickly this tactic became the norm.

Basic cyber weapons also became easier to build. Underground forums began offering “Malware-as-a-Service” kits with built-in AI that a novice could use to generate new malicious code or obfuscate existing malware for just a few dollars.

Additionally, AI helped attackers automate the grunt work of finding vulnerabilities. Machine learning systems can scan networks for weak points or even assist in writing exploit code. [The National Cyber Security Centre](#) globally warned that AI “will almost certainly continue to make elements of cyber-intrusion operations more effective and efficient,” including automating vulnerability research and exploit development.

Ultimately, AI became the cybercrime force-multiplier of 2025. It's important to note that many of these AI-driven attacks combined old tactics with new tech. Phishing, fraud, and malware are not new ideas, but AI supercharged their speed, scale, and effectiveness.

The key takeaway is that attackers don't need groundbreaking new hacks when they can achieve breakthrough efficiency with AI. [As one security expert put it](#), we have entered "a golden age of scammers" where AI lets every malicious email, call, or code snippet be precisely crafted to trick even vigilant targets. This shift has made it clear that relying on human users to spot telltale signs of a scam (like poor English or generic messaging) is no longer enough.

WHEN ATTACKS WENT AI-FAST, DEFENSE WENT AI-SMARTER

As AI-driven attacks surged in 2025, cybersecurity teams deployed AI as a strategic countermeasure. While adversaries used AI to scale threats, defenders leveraged it to accelerate detection, enhance intelligence, and speed high-precision response. The result marks a shift from playing catch-up to reclaiming advantage.

AI technology in security isn't new as ML (machine learning) has powered anomaly detection for years. However, 2025 marked the inflection point when AI matured into mainstream solutions, driving mission-critical operations.

Here's how AI redefined cybersecurity defense in 2025:

Early Detection Through AI "Anomaly" Sensors

One of AI's greatest advantages is finding patterns and anomalies across large data sets. In security, this means AI can learn what "normal" activity looks like for each user, device, and application in an organization, and flag subtle deviations that might indicate a threat.

In 2025, many organizations deployed AI-driven monitoring systems that watch network traffic, logins, file changes, and more — searching for abnormal behavior instead of known malware signatures. This approach, often called User and Entity Behavior Analytics (UEBA), dramatically improved detection of brand-new or stealthy attacks that traditional tools often miss. For example, [AI-based systems](#) identified unusual login patterns or data access behaviors that hinted at an insider threat or a hacker using stolen credentials, even when no known malware was involved.



In high-risk environments like banking, similar AI models achieved detection rates as high as 98% for certain attack types. By moving beyond signature-matching to behavior-based detection, defenders can catch novel attacks (like polymorphic malware or zero-day exploits) much earlier in the kill chain.

AI-Powered Threat Intelligence and Prediction

Another defensive improvement was using AI to proactively anticipate attacks. Threat intelligence feeds have long existed, but in 2025, they became supercharged with AI. ML platforms now ingest vast amounts of global threat data — from hacker chatter on the dark web to vulnerability disclosure trends — and analyze it to predict what's coming next.

For instance, AI can correlate hints from disparate sources and warn a company with “The type of vulnerability in your VPN device is likely to be exploited soon” or “We foresee a phishing campaign targeting your industry next quarter.” This predictive analytics capability helped some organizations patch or prepare defenses before a new wave of attacks hit. AI gave security teams a kind of radar — a forward-looking view of emerging risk — so they could harden systems proactively rather than responding after the fact.

Autonomous Response and SOC Automation

Speed is critical during a cyber incident, and AI enables much faster response times. [Security operations centers](#) (SOCs) typically drown in thousands of alerts a day, many of which are false alarms. AI stepped in as a “skills multiplier” for these teams. By automating alert triage and incident response, AI drastically reduced the workload and reaction time. For example, AI-driven security platforms can automatically correlate low-level alerts — stitching together a failed login here, a strange process there, and an odd data download — into one high-priority incident for analysts to investigate. This means human operators spend less time sifting noise and more time on real problems.

For certain well-understood threats, AI-enabled SOAR (Security Orchestration, Automation, and Response) systems took direct action without waiting for humans. If an endpoint was clearly infected with known ransomware, the system could isolate that machine from the network within seconds, or if a user's account showed a likely hijacking, it could automatically disable the account and require a reset.

Such autonomous responses helped contain incidents before they spread with significant payoff: companies that heavily adopted security AI and automation in their SOC reported millions of dollars less in breach costs on average, and incident lifecycles shortened by over two months on average.

Ultimately, AI-driven automation made cyber defense not only faster but also cheaper by preventing minor incidents from turning into major breaches.

AI Assistants for Security Analysts

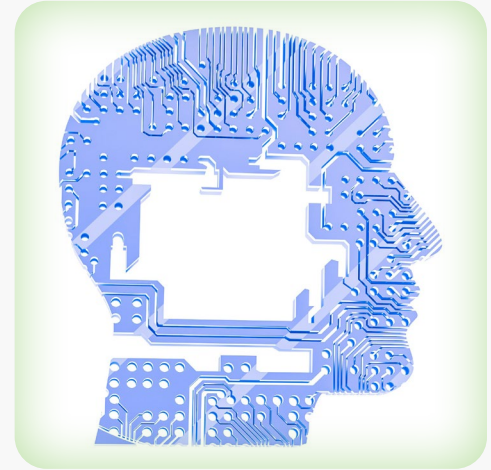
2025 introduced the era of the “[AI co-pilot](#)” for cybersecurity professionals. LLMs began integrating with security tools as natural language assistants. For example, an analyst could ask an AI assistant to summarize a flood of alerts or to explain the significance of a particular threat and get an instant, coherent answer.

These AI helpers pull together data from logs, past incidents, and threat intel to answer questions like “What unusual activity did we see on our database server last night?” or “Is this IP address associated with any known malware?” This significantly accelerates investigations and empowers less-experienced team members to work at a higher level.

Enhanced Filtering and Deepfake Detection

Traditional security tools quietly got smarter with AI under the hood. Email security gateways, for example, started using ML models to detect the subtle signs of AI-written phishing messages (identifying unusual linguistic patterns or grammar that is too-perfect in context). Some solutions also claimed to analyze email content with AI to catch threats that evade legacy spam filters.

On the identity side, new tools emerged to detect deepfakes. For instance, by analyzing audio calls for telltale digital artifacts or requiring “liveness” checks that are hard for deepfake videos to mimic (like asking a caller to turn their head or answer a personal question). Biometric logins that relied on face or voice had to be reconsidered, and many organizations shifted to more phishing-resistant multi-factor authentication solutions, such as physical security keys or one-time codes, instead of assuming a voice match was enough. In short, defenders in 2025 took a proactive approach to innovation.



AI became an indispensable part of cyber defense, enabling a shift from reactive to proactive security. With machine speed and ML on their side, security teams began evening the odds against a new generation of AI-empowered attackers.

AI CHANGED THE BATTLEFIELD 2025 WAS THE TURNING POINT

Beyond individual tools and tactics, 2025 delivered several AI-driven turning points for the cybersecurity industry. Together, they revealed how dramatically AI is reshaping the threat landscape, forcing organizations and regulators to reconsider how they respond.

Here are a few of the most notable trends and moments from the year:

The “Arms Race” Becomes Official

Industry leaders increasingly described cybersecurity in 2025 as an AI arms race. For the first time, both sides of virtually every incident had some level of AI involved, whether it was an attacker using an AI bot to craft an email, or a defender using an AI filter to flag it.

This dynamic was a hot topic at every major security conference. A dominant theme at the 2025 RSA Conference was “offensive AI vs. defensive AI,” and panels discussed scenarios like AI-driven malware fighting AI-driven detection in a continuous game of cat-and-mouse.

The general consensus: AI is now the central force driving the evolution of cyber threats and defenses.

Security budgets and R&D shifted heavily toward AI solutions, and companies that never used ML before were urgently evaluating AI vendors to mitigate risk and avoid falling behind.

Surge in AI-Era Attacks (and Awareness)

The volume and impact of attacks jumped sharply. We saw huge spikes in certain threat categories — phishing, BEC (Business Email Compromise), and fraud scams — directly tied to AI capabilities.

Phishing, in particular, reached record levels and overtook ransomware as the most-discussed threat of 2025. Law enforcement and agencies sounded alarms: the FBI issued official warnings that criminals are “leveraging AI” to dramatically increase the sophistication of phishing and fraud schemes. In public communications, deepfake scams went from a niche curiosity to a recognized menace.

Dark Web AI Services Emerge

On the attacker side, 2025 saw a commercialization of AI-driven crime tools. Early in the year, researchers noted the emergence of things like “WormGPT” and “FraudGPT” — essentially knock-off ChatGPT models with no ethical restrictions, sold to cybercriminals.

By mid-2025, these had evolved into more polished “dark LLMs” that bad actors could subscribe to for a fee. These underground AI platforms provided on-demand help for writing phishing content, creating fake documents, or even refining malware code. The accessibility of such tools meant that less-skilled criminals could launch more advanced attacks by outsourcing the “thinking” part to AI-as-a-service. This is a major shift as it broadens the pool of threat actors and increases the volume of attacks, because expertise is no longer a limiting factor. Nearly anyone willing to pay, or able to access a leaked model, can now leverage top-tier AI for nefarious ends.

In parallel, we also saw a niche industry of deepfake-as-a-service on criminal forums. Scammers who couldn't build deepfakes themselves could hire specialists offering to create fake videos or real-time impersonation for a price. This trend toward “AI outsourcing” in the criminal world served as a wake-up call for the cybersecurity industry, signaling the need to prepare for more attackers and a higher volume of attacks.

Major Investments and Products in AI Security

The cybersecurity market responded with a wave of innovation. 2025 saw dozens of startups and established vendors launching AI-driven security products.

The marketplace was flooded with AI branding: from AI-enhanced email security that claimed to detect AI phishing, and network monitoring tools with built-in ML analytics, to user authentication systems capable of spotting deepfake voices. AI is becoming a standard component of cybersecurity solutions. Analyst firm IDC predicted that by the end of 2025, 75% of security architectures will use AI or ML in some capacity. Another tangible shift comes with roles and responsibilities. Security operations job postings began listing “AI experience” or “familiarity with ML tools” as a desired skill, reflecting how security teams recognize the need for expertise in managing and interpreting AI-driven systems.



Regulatory and Ethical Scrutiny

With great power comes great responsibility and regulators in 2025 started grappling with the ramifications of AI in cybersecurity.

Governments moved to outlaw malicious deepfakes and require disclosures. Several jurisdictions introduced or passed laws making it a crime to create AI deepfakes for the purpose of fraud or electoral interference. Agencies like the U.S. Treasury's FinCEN flagged an increase in deepfake use in financial fraud reporting, which could lead to guidance for banks on verifying identities.

There was also recognition of the need for AI governance within organizations, not just to prevent misuse by criminals, but to ensure companies' own use of AI in security doesn't introduce new risks.

2025 proved that AI is not just another tool in the arsenal but instead, a paradigm shift. Cybersecurity incidents became faster, more complex, and more difficult to distinguish from legitimate activity due to AI's influence. But defenders also demonstrated that AI could significantly boost resilience when applied correctly.

The key lesson is that the fundamentals of security (protecting data, detecting intrusions, training users, etc.) remain the same, but the methods to achieve them must evolve. As we move into 2026, security leaders are taking stock of these lessons to adjust their strategies.

2026 STARTS NOW

AI DEFENSE BECOMES THE CISO MANDATE

In 2026, one thing is clear for CISOs and security teams: the old playbook needs an update. The AI-driven offense developments of 2025 demand changes in how organizations approach cybersecurity.

Here are some key takeaways and recommendations for adapting strategy, staffing, and tooling in the post-2025 era:

- 1. Assume AI, trust Nothing, and verify everything:** Make Zero Trust policy default. Validate high-risk requests through known, secondary channels to block AI impersonation.
- 2. AI defense is no longer optional:** Manual and signature-only security can't keep up. Prioritize AI-driven detection and automation in your 2026 security stack.
- 3. Watch behavior, not just thresholds:** Use ML to baseline normal activity and surface subtle anomalies across identities, logins, and data movement.
- 4. Secure the message, not just the link:** Modern email / content security must analyze language, sentiment, and media to detect AI-generated deception, not only malicious URLs.
- 5. Automate to contain, human-validate to decide:** Deploy SOAR / EDR / ID-response automation for rapid containment, with clear hand-offs to humans for escalation and judgment.

6. **Add friction for attackers, not users:** Implement lightweight identity assurance (PINs, call validation, audio / video screening) as deepfake defense to increase confidence without slowing the business.
7. **Govern AI like any high-risk third party:** Eliminate Shadow AI gaps; inventory AI tools, apply access controls, and assess risk before enterprise use.
8. **Train humans on AI fakes:** Awareness beats novelty. Regularly expose teams to real AI attack examples and enforce cross-channel verification training for high-risk roles.
9. **Prioritize AI fluency, not expertise:** Teams don't all need PhDs, just working knowledge of ML, training data, false positives / negatives, and adversarial AI risks to tune tools confidently.
10. **Add ML talent where it matters:** Hire or upskill 1 AI / ML specialist (e.g., security data scientist) to validate vendors, tune models, and tailor AI defense to your environment.
11. **Multiply analysts with AI, don't just hire more:** Use AI to handle tier-1 triage and repetitive tasks, speeding quality response, reducing burnout, and improving morale.
12. **Prepare for an ongoing battle and collaborate:** Finally, accept that the AI-cybersecurity arms race is ongoing. There is no "finish line" where one side wins. New AI models will emerge, attackers will find new exploits, and defenders will develop countermeasures, in a continual cycle. Build a strategy of continuous improvement and adaptation.

This can include:

- 🕒 **Regularly schedule red-team exercises or penetration tests** that specifically simulate AI-driven attacks (like deepfake phishing or AI-written malware) to see how your people and systems hold up.
- 🕒 **Participate in information-sharing** communities about AI threats. In 2025, many industry groups, and even vendors, started sharing insights on prompt injection attacks, deepfake indicators, and other AI-related threat intelligence. Tap into these communities to help keep you ahead of the curve.
- 🕒 **Keep an eye on regulatory developments** and be ready to comply with new AI requirements. It's better to build good governance now than be caught off guard by a law that, for example, mandates disclosure of AI-generated content or requires auditing your AI models for bias / security.

Heading into 2026, organizations that internalize these lessons will be far better positioned to handle the next wave of attacks. The overarching theme is resilience through agility — using AI to defend, training your people, and adjusting processes to mitigate AI-enhanced threats. Cybersecurity has always been about adapting to change and AI is just the latest, albeit powerful, change.

CONCLUSION

2025 showed us what the future of cybersecurity will look like. Generative AI has irreversibly changed the threat landscape, tearing down the old limitations of social engineering and malware creation. In turn, it has pushed defenders to reinvent how we protect systems and data.

The cat-and-mouse game between attackers and security teams is now turbocharged by algorithms on both sides. While this prospect can sound daunting, the experience of 2025 also offers hope: organizations that harness AI for defense and proactively address new threats can still stay one step ahead.

The key is to not underestimate the pace of change. Stay vigilant, stay informed, and embrace the tools and practices that will help you thrive in the age of AI-enhanced attack and defense.

ABOUT US

With more than 30 years of experience partnering with global enterprises on complex business and IT initiatives, SDG is a trusted provider of advisory, transformation, and managed services. The firm empowers organizations to strengthen cyber resilience by integrating AI into identity, threat, and risk management solutions that protect digital assets and deliver measurable business value.

Learn more at www.sdgc.com.



■ 75 North Water Street
Norwalk, CT 06854
■ 203.866.8886
■ sdgc.com

Contact Us: solutions@sdgc.com