COFENSE

AI and the Future of Phishing Defense

Key Insights and Predictions for 2026

INTRODUCTION

Email remains the primary communication channel for most organizations, but it is also one of the most frequently exploited gateways for cyberattacks. As threat actors accelerate their use of artificial intelligence, phishing campaigns are becoming more targeted, more convincing, and harder to detect. At the same time, AI is transforming the phishing-defense landscape itself, reshaping how security teams identify, analyze, and respond to threats. Entering 2026, this rapid evolution raises the stakes for organizations working to stay ahead of increasingly adaptive adversaries.

To help bolster your security posture, we've gathered predictions from Marc Olesen, CEO, alongside insights from Cofense experts, to highlight the trends redefining email security and the vulnerabilities organizations must prepare for in the coming year. Whether you're a seasoned security professional or managing cybersecurity for your organization for the first time, these perspectives will help you navigate an environment where AI is driving both the attacks and the defenses of the future.

Strengthening phishing defense in 2026 will require cybersecurity leaders to think strategically to stay ahead of advanced AI-driven threats. As attackers refine their techniques and uncover new vulnerabilities, organizations must be prepared for the challenges that lie ahead. Cofense experts have identified the key challenges security teams must anticipate and address to maintain a strong and resilient security posture in the year ahead.

MARC OLESEN
Chief Executive Officer

Human–AI partnership will define cyber resilience

Organizations with the strongest resilience and security will balance automation efficiency with human intelligence. Many believe that AI will be better, faster, and cheaper, but accepting that AI comes with limitations will be vital and cannot replace the accuracy and contextual understanding of human intelligence. 2026 will be about amplifying human resources for what they are best at - analytical accuracy, strategic agility, and critical thinking - and combining that with the best AI capability to enable fast, efficient triaging, and prioritization.

WHAT CAN YOU DO TO PREPARE?

- **Create an AI-augmented Security Operations Center (SOC)**
  Re-engineer SOC workflows so AI handles volume and speed to clear the noise - triaging low-level known threats, pattern matching, and prioritization - while analysts focus on high-impact, high-risk and unknown threat verification and decisions. Ensure SOC teams and leadership understand AI logic, limits, and outputs to develop this workflow and apply resources accordingly. This will help build a picture of an organization's unique environment while identifying potential gaps for improvement.
- **Invest in a human-supervised culture that feeds AI**
  Teams that see AI automation as an enabling tool rather than a threat will help embrace and drive continuous improvement. When the human layer identifies a new threat, feedback systems should be in place to injest new indicators to be catalogued, shared, and programmed into tools so models evolve with real-world intelligence.

MARC OLESEN
Chief Executive Officer

AI transparency and governance will become a hot topic

The “black box” nature of AI is becoming a growing concern and poses a big risk in cybersecurity. As organizations continue to adopt AI tools in their security stack, the importance of being able to demonstrate, justify, and audit decisions and actions will likely become a priority for governing bodies. How we regulate these tools moving forward to ensure compliance with the likes of GDPR, DORA, HIPAA and many more, will become a strong topic of conversation. It will be up to leadership to drive best practices ensuring transparent, ethical, and compliant AI use, while regulators work out how best to set controls to protect data privacy.

WHAT CAN YOU DO TO PREPARE?

- **Implement explainable AI and auditable controls**
  Demand transparency from vendors on how models make decisions and use personal data, and better yet, ensure control over specifically what and when AI is being used. Deploy solutions with built-in audit trails that allow teams to trace why a particular threat was flagged, or missed, across the security stack. Implementing these best practices will make future compliance and adjustments for new regulations much easier.
- **Establish AI governance at the board level**
  Make AI and automation a standing item on the board agenda, and ensure leadership is trained on AI’s capabilities, limitations, and ethical implications. Establish clear accountability frameworks to maintain transparency and compliance in AI-driven decision-making. Taking these steps now will help prevent or limit future reputational and financial risk.

JOSH BARTOLOMIE
VP, Global Threat Service

Organizations will experience an automation paradox in security operations

Accelerating AI adoption within security operations while simultaneously reducing analyst headcount will create a critical blind spot in 2026. As threat actors weaponize generative AI to create polymorphic phishing campaigns at scale, the pattern-recognition limitations of current AI detection systems will become starkly apparent. This convergence of leaner SOC teams paired with over-tuned automation will result in a measurable increase in successful novel phishing attacks that bypass traditional and AI-augmented detection mechanisms.

WHAT CAN YOU DO TO PREPARE?

- Invest in analyst expertise, not just tooling and automation. Implement “human-in-the-loop” mandates for all AI-driven security decisions affecting critical assets or novel threat patterns, with clear escalation thresholds that trigger expert analysis. By prioritizing retention and training of senior analysts who can identify subtle anomalies that AI systems miss, these practitioners will become force multipliers for your automation.
- Build redundancy into AI threat detection logic. Deploy multiple AI models with different training sets alongside traditional rule-based systems. No single detection method should be a point of failure.

JOSH BARTOLOMIE

VP, Global Threat Service

The timeline from phishing campaign launch to compromise will compress dramatically

AI will allow threat actors to rapidly test, refine, and scale phishing campaigns before traditional threat intelligence or detection systems can react. Organizations used to having hours or days to respond will face compromises within the first few hours of an attack. Breaches that were once contained will escalate as attackers use AI to weaponize stolen data—driving faster, more sophisticated credential stuffing, follow-on phishing, and account takeover attempts. What previously took weeks to monetize will occur in days, creating "breach cascade" scenarios where one compromise fuels additional AI-driven attacks across an organization's ecosystem.

WHAT CAN YOU DO TO PREPARE?

- **Integrate real-time, phishing-focused threat intelligence and strengthen human-driven detection.** Continuously update your security stack with actionable intelligence on active campaigns, attacker infrastructure, and evolving TTPs so threats are identified by patterns—not static signatures. At the same time, empower and train employees to report suspicious emails and treat user-submitted signals as mission-critical data. Humans can spot intent, social engineering cues, and contextual anomalies that automated systems miss, making an engaged workforce your most adaptive defense against AI-generated, polymorphic phishing variants.
- **Build layered detection that combines IOC accuracy with campaign and TTP context.** Use real-time phishing IOCs for fast blocking of known threats, but pair them with campaign-level intelligence to spot consistent attacker behaviors, infrastructure patterns, and TTPs that persist across polymorphic variants. This dual approach enables rapid prevention when signatures exist while still detecting novel and zero-day phishing attacks that lack identifiable IOCs—maximizing coverage against both known and emerging threats.

MAX GANNON
Manager, Intelligence Analysis

Context becomes the only defense against modern phishing

Traditional email security controls that rely on scanning embedded URLs will increasingly fall short. Threat actors are escalating their use of redirection tactics such as legitimate open redirects, link shorteners, and trusted platforms like Microsoft Forms to obscure malicious destinations. These techniques, combined with convincing brand impersonation and AI-generated phishing lures, routinely bypass secure email gateways (SEGs) and AI-based filters. Worse, the network traffic and file signatures involved often appear benign, creating a dangerous blind spot for automated defenses. As seen in recent Remote Access Trojan and Remote Access Tool campaigns, even network defenders often dismiss these signs as legitimate IT activity, highlighting how context, not content, makes the difference.

WHAT CAN YOU DO TO PREPARE?

- **Adopt context-aware threat detection:** Move beyond static, content-based scanning by implementing defenses that evaluate behavior, environment, and intent—not just URLs or message elements.
- **Empower users as critical sensors:** Provide training that teaches employees to recognize fake login pages, identify suspicious activity even after engagement, and confidently report anything that seems off, strengthening your human-enabled security layer.

MAX GANNON
Manager, Intelligence Analysis

Yesterday's bugs will become tomorrow's breach headlines

Attackers will increasingly use offensive AI to uncover unpatched variations of known vulnerabilities, focusing on systems where updates only partially resolved the issue. Instead of relying on undiscovered zero-days, threat actors will revisit old CVEs, using AI to explore slightly altered versions that slip past incomplete fixes or less-tested environments. This technique allows them to scale attacks quickly while staying just outside the scope of existing detections.

This same tactic will be applied to email security tools. If one phishing email manages to bypass a filter, AI will be used to rapidly generate dozens of near-identical variants, tweaking elements like formatting, sender behavior, or content structure until a new successful bypass is found. These small adjustments can be enough to confuse pattern-based and AI-driven defenses, making fast, repeated exploitation more likely.

WHAT CAN YOU DO TO PREPARE?

- **Reinforce vulnerability management:** Regularly validate that patches are fully applied, test for variant exposures, and prioritize monitoring for slight deviations of previously resolved CVEs.
- **Harden email defenses against rapid iteration:** Deploy layered detection capable of identifying behavioral patterns, not just content, and ensure security teams can quickly analyze and block AI-generated phishing variants.

CHANCE CALDWELL
Sr. Director, Phishing Defense Center

AI Will Turn Oversharing into Phishing Bait

Generative AI will enable threat actors to create tailored phishing campaigns at scale, using publicly available data like LinkedIn profiles, org charts, and even social media posts. With AI's ability to craft realistic emails, mimic writing styles, and automate iterations, phishing attacks will become nearly indistinguishable from legitimate messages, increasing the risk for every organization.

WHAT CAN YOU DO TO PREPARE?

- **Reduce your public footprint:** Audit and limit the information employees and leaders share online, and train staff to understand how publicly available details can be weaponized in targeted phishing attacks.
- **Reinforce fundamental security controls:** Strengthen multi-factor authentication, domain monitoring, and email warning systems, and provide security awareness training that helps users recognize sophisticated, AI-crafted phishing, not just obvious red flags.

CHANCE CALDWELL
Sr. Director, Phishing Defense Center

Remote Access Abuse Will Increase in 2026

Threat actors increasingly rely on tools like ConnectWise and AnyDesk, blending into normal IT activity to deliver malware and avoid raising alarms. These tools eliminate the need for custom infrastructure, making attacks faster, cheaper, and harder to spot. This trend dominated 2025 and shows no signs of slowing down in 2026.

WHAT CAN YOU DO TO PREPARE?

- **Tighten control of remote access tools:** Limit the use of remote tools, enforce strict access permissions, and continuously monitor for unusual or unauthorized activity to distinguish normal behavior from abuse.
- **Train users to recognize remote-access phishing tactics:** Provide training that helps employees identify attempts to deploy these tools through fake tech support, document lures, or other social engineering methods, strengthening early detection.

CONCLUSION

The predictions from our experts make one thing clear: phishing defense in 2026 will be defined not only by rapidly advancing AI-driven threats, but by how effectively organizations balance AI capabilities with human expertise.

Staying ahead requires more than new tools—it demands strategic thinking, continuous human-AI learning, strong governance, and empowered users who can act as critical sensors. With attackers blending into normal activity, abusing legitimate tools, and scaling highly tailored phishing campaigns, resilience will depend on combining intelligent automation with human insight and vigilance.

Want deeper insight into these predictions and actionable steps to safeguard your organization in the year ahead? Watch our 2026 Email Security Predictions webinar on-demand!

About Cofense

Cofense® provides the world's most effective business email threat detection and remediation solutions. The Cofense PhishMe® Email Security Awareness Training (SAT) Platform with Risk Validation and the Cofense Phishing Threat Detection and Response (PDR) Platform are powered by over 35 million Cofense-trained employees who report phishing and other dangerous email threats in real time. Exclusive to Cofense, our network detects and eradicates threats other email security systems miss and removes them from our customers' inboxes. For more information, visit cofense.com or connect with Cofense on X and LinkedIn.