THE RISE OF AI

A NEW ERA OF PHISHING THREATS

PHISHME
COFENSE

Introduction

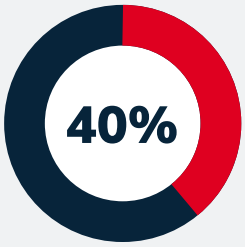
The rapid rise of artificial intelligence (AI) has revolutionized industries and transformed the digital landscape. While AI offers immense potential for innovation and efficiency, it also opens the door for new challenges and vulnerabilities. One critical area where the influence of AI is creating complex obstacles is in cybersecurity, particularly email security. From sophisticated, never-before-seen phishing schemes to the automation of malware delivery and polymorphic attacks, AI is reshaping how organizations need to approach and address threats that bypass perimeter defenses and land in employee inboxes.

Email remains the #1 threat vector **and its extensive business usage makes it a prime target for threat actors looking to exploit its vulnerabilities.**

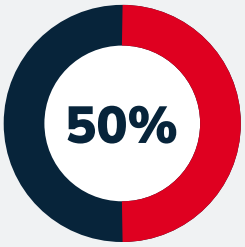
With the advent of AI-driven advancements, attackers now have access to unparalleled tools that allow them to amplify the scale and effectiveness of cyber threats. While AI empowers cybersecurity experts to create advanced tools that detect and respond to threats with speed and precision, cybercriminals are rapidly evolving their tactics to exploit these same technologies. Offensive AI will always maintain an edge over defensive AI, as it operates without the legal and ethical constraints that safeguard responsible development of AI. Threat actors can train their models on unethically sourced data and use distributed networks of compromised computers to run processing-intensive algorithms that would be cost-prohibitive for legitimate organizations.

This lack of oversight, combined with the ability to innovate without restrictions, allows offensive AI to evolve faster, outpacing the majority of defensive tools.

Cofense analysts see the steady rise of AI-generated phishing emails every day, observing firsthand how threat actors continually refine their methods to deceive and exploit targets. Leveraging their deep expertise, our experts have meticulously analyzed the tactics and techniques behind these increasingly deceptive campaigns, uncovering the strategies that make them so effective. Starting with data from early 2024, they analyzed key trends and uncovered actionable insights, all outlined in this report to help organizations protect their assets in this dynamic era of AI.



More than 40% of malware families detected in 2024 were new to Cofense.

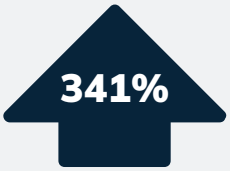


Of the newly created malware families, **almost 50% were Remote Access Trojans.**

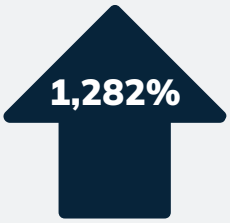
The Cofense Phishing Defense Center saw, on average, **1 malicious email every 42 seconds throughout 2024.**

"Business Email Compromise" emails **increased 70% from 2023.**

Industries with the largest increase in reported malicious emails:



Education: **341% increase in reported malicious emails.**



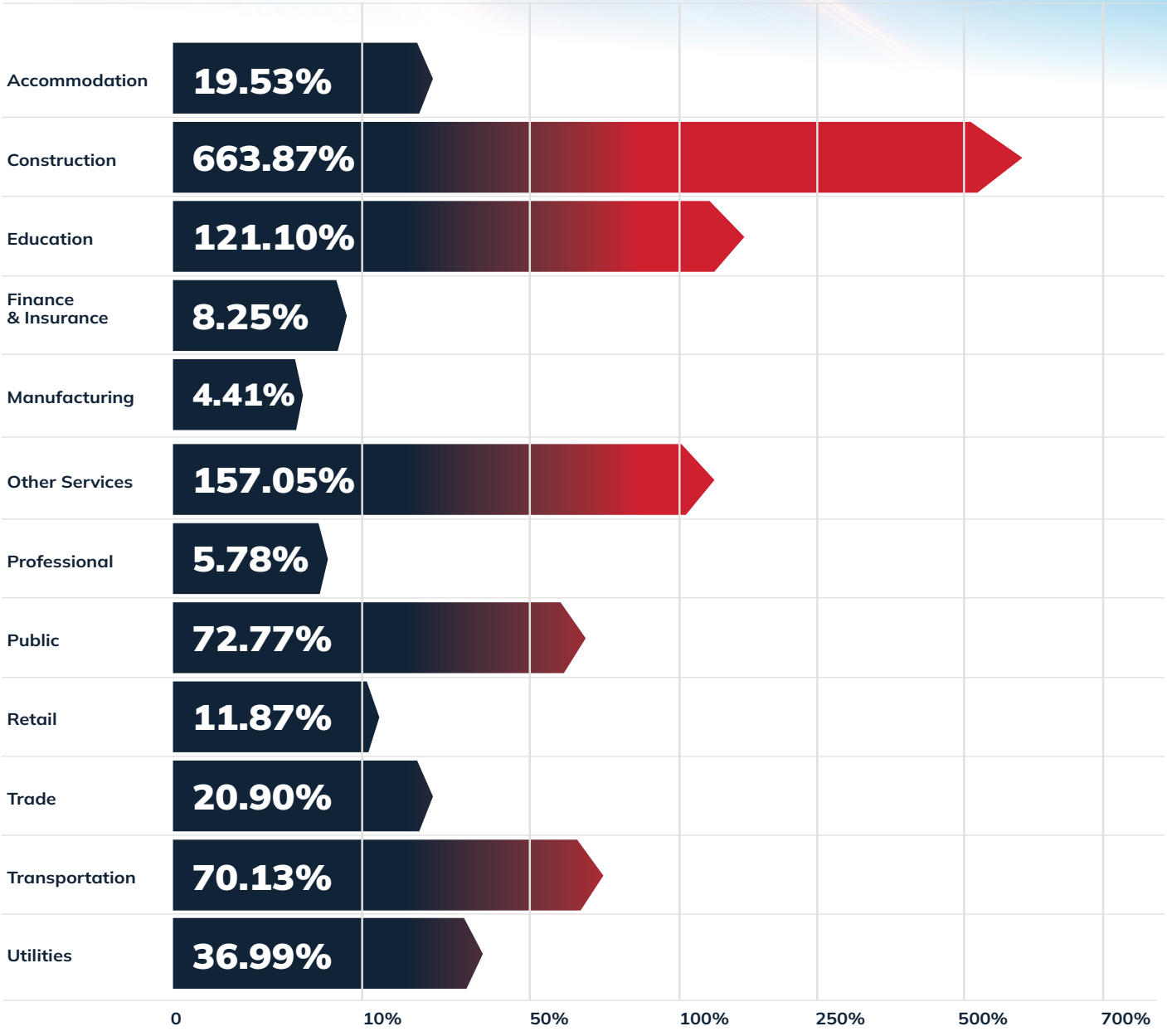
Construction: **1,282% increase in reported malicious emails.**

See page 4 for a complete breakdown by industry.

Taxes-related campaigns **increased by 340%.**

Campaigns utilizing legitimate files **increased by 575%.**

Percentage Change in Malicious Emails Bypassing Email Security Perimeters Year Over Year



Understanding the breakdown of phishing threat trends by industry is critical to enhance your organization's cybersecurity posture. This data reveals how threats are evolving and highlights potential weaknesses in current email security measures. By analyzing these trends, organizations can refine their defenses, allocate resources effectively, and stay ahead of emerging threats, minimizing the risk of a successful breach.

Top Trends To Watch

1. AI is accelerating the creation and spread of new malware families by automating code generation, personalizing attacks through advanced data analysis, and evading detection with adaptive learning algorithms.

2. Threat actors are leveraging generative AI to craft highly targeted and cosmetically perfect campaigns.

3. Business email compromise phishing is increasing due to the use of AI to automate email generation, adjust strategy based on target responses, and evade spam filters by using slight variations in text or structure to avoid triggering alarms.

4. The rise of polymorphic phishing attacks seeds the need for combined AI + human-vetted defense.

5. The diversification of phishing email content, structures, narratives, and technical nuances has greatly increased.

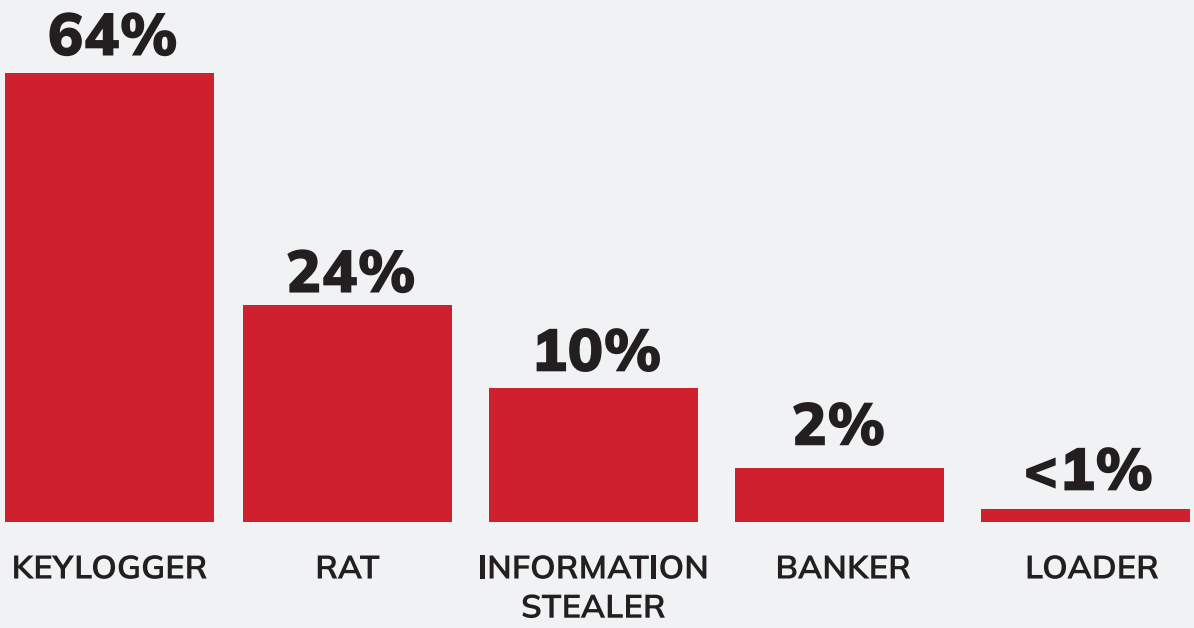
Trend #1:
AI Accelerates Malware

Generative AI has enabled cybercriminals to accelerate the creation and spread of new malware families. Over 40% of the malware families analyzed by the Cofense Intelligence team in 2024 were new compared to the data from 2023. By automating code generation, AI streamlines the development of sophisticated malware, reducing the technical expertise required by attackers. Automated tools powered by AI can efficiently produce variants of existing malware, ensuring an almost endless supply of harmful programs that are capable of exploiting vulnerabilities at an unprecedented rate.

"Threat actors are creating, manually or with the help of AI, large numbers of new malware families, specifically focusing on Remote Access Trojans (RATs) that have additional capabilities. Although keyloggers were the dominant malware type in 2024, the new malware families being predominantly RATs indicates a shift in future trends."

-Max Gannon, Mgr, Intel Analysis

Top 5 Malware Types:



Of the newly created malware families in 2024, almost 50% were Remote Access Trojans, with that number increasing in 2025.

AI also enables malware to feature adaptive learning algorithms, which allow it to evolve and respond to security defenses in real-time. These algorithms can analyze patterns in cybersecurity protocols and modify the malware's behavior to sidestep detection mechanisms, such as antivirus software or intrusion detection systems. This adaptability makes AI-driven malware incredibly challenging to pinpoint and neutralize, as its tactics and techniques constantly shift to stay ahead of defensive measures. Consequently, traditional reactive responses to cyber threats are proving to be less effective against these advanced forms of digital attacks.

Top 5 Malware Family Characteristics

Name	Information Stealing	Keylogging	Remote Access	Loader Capabilities	Backdoor Controls
Agent Tesla Keylogger	✓	✓	✓		
Remcos RAT	✓	✓	✓	✓	✓
FormBook	✓	✓			
Snake Keylogger	✓	✓			
XWorm RAT	✓	✓	✓	✓	✓

The integration of AI into malicious activities poses an ongoing risk for security teams. While AI continues to accelerate the evolution of malware, it is critical to invest in solutions that can predict, detect, and mitigate these threats proactively. Pairing cutting-edge technology with human-vetted intelligence ensures a deeper understanding of evolving attack patterns and tactics. This combination not only enhances detection accuracy but also strengthens response capabilities because understanding the context of threats is the key to effective mitigation.

Trend #2: Generative AI Targeting Perfection

Threat actors are leveraging generative AI to craft highly targeted and cosmetically perfect campaigns.

The rise of generative AI tools has introduced a new era of customization in cyberattacks, enabling threat actors to craft highly targeted and cosmetically flawless campaigns at scale with minimal effort. By leveraging advanced algorithms, malicious actors can generate phishing emails, fake websites, and deceptive social media profiles that appear highly credible to unsuspecting victims.

By analyzing publicly available data, such as company names and job titles, from social media platforms, leaked databases, and online footprints, cybercriminals can create customized messages that resonate with specific targets. For instance, an AI-generated phishing email might reference a victim's recent purchases, professional affiliations, or interests, thereby increasing the likelihood of engagement. This level of precision, combined with the natural language capabilities of generative AI models, results in attacks that are both highly convincing and alarmingly effective.

> **Generative AI will multiply losses from deepfakes and other attacks 32% to 40 billion annually by 2027.**
> - DELOITTE

Since threat actors no longer need to spend extensive time and resources designing social engineering schemes manually, they can quickly and seamlessly launch highly coordinated and multi-dimensional campaigns. This automation allows for a higher frequency of attacks, putting additional strain on existing cybersecurity measures and response teams, which often struggle to keep pace with the sheer volume of threats.

Trend #3: AI Is Increasing Effectiveness of Business Email Compromise

Business email compromise (BEC) attacks are increasing due to the use of AI to automate email generation, adjust strategy based on target responses, and evade spam filters by using slight variations in text or structure to avoid triggering alarms.

In 2024, the Cofense Phishing Defense Center identified a new BEC strategy that exploded in popularity, generating a lot of confusion and trouble for target organizations. This tactic involves BEC emails in which the threat actor sends what appears to be a forwarded email chain to an employee, asking them to take care of an overdue invoice. The email chain will "show" a high-ranking member of the victim's organization, such as a CEO, approving an invoice and instructing the victim to ensure it is processed promptly. In actuality, the CEO has not communicated with the organization in question. These emails most commonly show the original request coming from an "@consultant.com" address, which can be used as an indication of potential malicious activity.

“Business Email Compromise” attacks increased 70% from 2023 to 2024.

Trend #4: Polymorphic Phishing on the Rise

The rise of polymorphic phishing attacks seeds the need for combined AI + human-vetted defense.

The analysts in the Cofense Phishing Defense Center observed a sharp rise in polymorphic phishing attacks in 2024, making it increasingly difficult for organizations to detect and mitigate threats. Unlike traditional phishing methods, polymorphic phishing attacks rely on dynamic changes to the appearance and structure of malicious emails or links. Attackers use sophisticated algorithms to alter subject lines, sender addresses, and email content in real time, effectively bypassing static signature-based email filters. Each iteration of the phishing attempt is uniquely crafted, reducing the probability of detection and enabling threat actors to execute their campaigns with alarming efficiency.

AI-powered systems are adept at analyzing vast amounts of email traffic and detecting subtle patterns indicative of malicious activity. They can identify anomalies such as unusual sender behaviors, suspicious domains, or irregularities in message structures. However, AI's strength is not without its limitations. Sophisticated polymorphic attacks, designed to evade algorithms by constantly altering their traits, can still slip through automated defenses. This is where the human element becomes critical, as skilled analysts can offer contextual understanding and intuition beyond the reach of algorithms.

> **"Polymorphic phishing attacks are difficult to find and remediate with traditional security tools due to the lack of commonality between each phishing attack. It takes humans, who are able to use a wider scope of context, to identify and manage these emails as they bypass AI security technology with new techniques."**

- Chance Caldwell, Sr. Director, Phishing Defense Center

Trend #5: Phishing Diversification

The diversification of phishing email content, structures, narratives, and technical nuances has greatly increased.

Cybercriminals have shifted their focus from increasing the number of phishing emails to diversifying their strategies. This diversification has resulted in a much larger variety of content, making phishing campaigns harder to detect and preventing recipients from becoming accustomed to recurring patterns.

Taxes-related campaigns increased by 340%

Multi-factor, authentication-related campaigns decreased by 64%

Campaigns utilizing steganography increased by 37%

Campaigns utilizing legitimate files increased by 575%

Cybercriminals increasingly exploit a broader range of emotional triggers, leverage current events, and mimic trusted brands to manipulate recipients into taking action. Whether it's pretending to be a well-timed tax subsidy offer, exploiting the fear of account suspension, disguising messages as urgent charity appeals during natural disasters, or imitating trusted brands with enticing offers, these schemes are expertly crafted to exploit human emotions and vulnerabilities through precise psychological manipulation. The adaptability and creativity of these narratives further erode the effectiveness of historical defense mechanisms such as pattern recognition.

Microsoft has been identified by Cofense Intelligence as the most frequently spoofed brand in 2024.

Conclusion

The increasing use of AI by threat actors has ushered in a new era of phishing attacks. AI is enabling threat actors to craft emails, SMS texts, deepfake videos, and audio content that are nearly indistinguishable from legitimate communications. This technology drastically reduces the time and cost of creating email-based attacks while increasing their effectiveness, leaving employees and organizations more vulnerable to data theft, financial loss, and reputational damage. And it's just the beginning – with time, AI-based attacks will only become more personalized and deceptive, leaving organizations at an increased risk of falling victim to threats that exploit both technological vulnerabilities and human error.

Cofense is uniquely positioned to help organizations mitigate the risks of AI-enabled phishing through an integrated approach that combines human-supervised AI with cutting-edge technology. By leveraging a global network of over 35 million trained end users, our solutions can detect and analyze malicious emails in real-time, ensuring organizations stay one step ahead of attackers. Additionally, Cofense emphasizes the importance of employee education, arming teams with the knowledge and skills needed to identify and report suspicious emails. This combination of technology and human insight strengthens detection capabilities, empowers employees as an active line of defense, and fosters a security culture that protects organizations from AI-driven threats with speed and precision.

Contributors

Josh Bartolomie
Chief Security Officer;
VP, Global Threat Services

Chance Caldwell
Sr. Director,
Phishing Defense Center

Max Gannon
Mgr, Intel Analysis

Rachel Roldan
VP, Product

PHISHME
COFENSE

Cofense® is the original and leading provider of security awareness training and phishing simulation, offering global enterprise-level advanced email threat detection and remediation solutions. Cofense PhishMe® and Cofense Phishing Detection and Response (PDR) offer the world's only platforms to leverage over 35 million Cofense-trained employees who actively report suspected phishing and other dangerous email threats in real-time. Exclusive only to Cofense, this reporting system ingests and catalogs thousands of threats per day that are missed by current email gateway technologies and then eradicates those threats from customer inboxes. In short, Cofense sees and stops threats other email security systems miss. Please visit www.cofense.com or connect with us on X and LinkedIn for additional information.