CYBER
THREAT
ANALYSIS

Recorded Future®

By Insikt Group®

January 28, 2025

2024 Annual Report

The most significant attacks of 2024 demonstrated how SaaS application proliferation amplified the effect of stolen credentials. Lack of MFA and other misconfigurations contributed to threat actors accessing sensitive data.

Criminal groups proliferated in the wake of law enforcement action. The number of ransomware attacks remained similar to last year, though the number of families and variants increased.

Generative AI helped nation-states scale up influence operations. As voters in over 70 countries went to the polls, adversarial influence operations used generative AI tools to amplify inauthentic content.

Foreword

Our annual threat reports provide an in-depth look at threat actors' key tactics, techniques, and procedures (TTPs) and motivations from the past year. These inform anticipated TTPs for the year ahead. You can use this information to better detect threats and manage cyber risk.

Learn from the adversary's 2024 playbook to strengthen your security posture.

This report presents the industry's most comprehensive intelligence analysis from 2024, a year in which enterprise attack surfaces continued to expand, cybercriminal activity carried on despite high-profile law enforcement actions, and state-sponsored threat actors targeted critical infrastructure and elections.

Read on for details related to key threat actors, their targets, and the TTPs they used so you can identify and address security control gaps in your systems. You'll learn how:

- Threat actors exploited software-as-a-service (SaaS) setups and valid credentials, notably in the ALPHV and RansomHub ransomware attacks on Change Healthcare and the UNC5537 attack on Snowflake.
- Despite law enforcement disruptions of major ransomware-as-a-service (RaaS) groups like LockBit and ALPHV and the fact that ransomware activity remained consistent year over year, the number of new ransomware groups increased.
- As over 2 billion voters went to the polls worldwide, China, Russia, and Iran used AI-generated content in an attempt to mold public opinion as part of malign influence operations.

Prevent future attacks with a roadmap for 2025 and beyond.

Review our predictions regarding memory-safe coding, cryptocurrency fraud, third-party risk management, and more. Use the information to craft a threat intelligence strategy that can help safeguard your data, brand reputation, and bottom line. Our predictions include:

- Developers will use AI to accelerate the transition to safer code.
- Cryptocurrency fraud will lead to a market-destabilizing event.
- A major breach will very likely result from the implementation of GenAI into enterprise workflows, or the abuse of AI for effective impersonation.
- More companies will report Chinese pre-positioning activity, demonstrating China's capability to conduct disruptive operations across a broader range of critical industries.

On behalf of Recorded Future's Insikt Group, I hope this report will help you stay one step ahead of the adversary and avoid disruption to your business in 2025.

Levi Gundert
Chief Security & Intelligence Officer

Executive Summary

Two parallel trends shaped the cybersecurity landscape in 2024: the resilience of criminal networks in the face of increased law enforcement disruptions and the growing complexity of enterprise attack surfaces. The cybercriminal marketplace has demonstrated resilience, as criminals adapted to withstand persistent law enforcement takedowns through reorganization, while enterprise networks have become almost too complex as more essential business processes shift to third-party, cloud-based applications.

Multiple law enforcement disruption efforts targeted ransomware operations (such as LockBit), denial-of-service operators, and malware-as-a-service (MaaS) providers throughout 2024. While there was a temporary decrease in ransomware victims immediately after law enforcement seizures or action, overall volumes for reported attacks have not decreased significantly. This adaptability demonstrates how the service-based criminal ecosystem enables threat actors to pivot between providers quickly while maintaining operational continuity. Moreover, the fundamental market dynamics driving cybercrime remain strong, motivating criminals to adapt despite setbacks.

On the other hand, the average enterprise now employs 371 distinct software-as-a-service (SaaS) products, creating a larger attack surface that has become increasingly challenging to defend effectively. Credential compromise for SaaS products and environments have led to significant breaches, as evidenced by the incidents affecting Change Healthcare and Snowflake. Critical vulnerabilities in widely deployed enterprise products, such as Ivanti equipment and Confluence, have also resulted in widespread security breaches.

At the same time, state-sponsored threat actors and their hacktivist proxies (which were primarily linked to China and Russia) launched disruptive attacks on critical infrastructure networks such as water treatment facilities to advance their geopolitical objectives. In addition to cyberattacks, China, Russia, and Iran used generative AI to conduct information operations and spread inauthentic political content around the world. These threat actors also took advantage of heightened complexity — both in industrial technology and in the information ecosystem — to obscure evidence of government involvement in these activities.

In response to an entrenched cybercrime industry, an unwieldy number of third-party considerations, and increasingly emboldened state-sponsored threat actors, organizations must move beyond traditional perimeter-based defenses to adopt more robust, scalable, and automated security architectures. These security solutions need to match the complexity of enterprise operating environments and the proliferation of increasingly advanced threat actor toolsets. That said, based on major breaches like Snowflake in 2024, organizations must first prioritize straightforward and effective solutions — like multi-factor authentication (MFA) — over newly complex security tooling.

Key Findings

- **The growing adoption of SaaS applications made identity exploits more effective.**
  The most significant attacks of 2024 demonstrated how SaaS application proliferation in corporate environments amplified the effect of credentials stolen via infostealer malware.
- **Criminal groups proliferated in the wake of law enforcement actions.**
  Law enforcement successfully disrupted major ransomware-as-a-service operators, including LockBit. However, distinct ransomware families and variants increased sharply in late 2024 as criminals used exposed source code and builders to launch independent operations.
- **Disruption and data made cybercrime pay.**
  Manufacturing and health care remained the most targeted industries by ransomware and extortion operators in 2024, demonstrating the continued payoff for cybercriminals to threaten real-world disruption. Meanwhile, the number of databases for sale on criminal forums increased by 20% over the last year, with telecommunications, healthcare, and education databases commanding the highest prices.
- **Increasing global hostilities led to critical infrastructure disruption.**
  State-sponsored threat actors and their hacktivist proxies weaponized critical infrastructure threats to advance geopolitical goals. China-linked threat actors were identified pre-positioning on critical networks, while Russia- and Iran-backed hacktivists targeted water treatment facilities for maximum visibility.
- **Nation-states took advantage of generative AI to level up influence operations.**
  As voters in over 70 countries went to the polls, influence operations from China, Russia, and Iran used generative AI tools to expand content production and reach. All three nations continue investing in GenAI research.
- **Tactics, techniques, and procedures emphasized defense evasion.**
  Tactics once limited to sophisticated, state-resourced threat actors are becoming increasingly common in criminal operations. These include using legitimate tools to evade detection and developing malware in Go and Rust coding languages.

Table of Contents

Growing Adoption of SaaS Leads to Identity Exploits 4
- **SaaS Applications Provide New Opportunities to Exploit Stolen Credentials** 4
- **Infostealer Infections Increasingly Target Personal Devices and Obtain More Credentials per Infection, Increasing the Risk to SaaS Ecosystems** 5
  - *Geographic Dispersion of Workforces Increases Insider Threat Risks* 5
  - *Looking Ahead: Taking Steps to Secure Cloud Identities* 6

Extortion Groups Proliferate Despite Law Enforcement Action 6
- **Law Enforcement Actions Disrupt Two Major Affiliates** 6
  - *Looking Ahead: Adapting to Variety in the Ransomware Ecosystem* 9

Industry Analysis Gives Insight into Threat Actor Priorities 10
- **Threat Actors Seek to Amplify Urgency to Increase Extortion Payouts** 10
- **Value of Stolen Data Driven by Future Monetization** 11
  - *Looking Ahead: Anticipating the Highest "Return on Infection"* 12

Escalating Global Hostilities Drive Critical Infrastructure Targeting 13
- **Chinese State-Sponsored Pre-Positioning Demonstrates Capabilities to Disrupt US Critical Infrastructure** 13
  - *Salt Typhoon's Massive Telecom Hack and US-China Relations* 13
- **Russian-Linked Sabotage Efforts Indirectly Advance War Aims** 14
  - *Looking Ahead: Geopolitical and Cyber Convergence* 15

Generative AI Accelerates Spread of Inauthentic Content in Historic Election Year 15
- **State-Sponsored Adversaries Continue Experimenting with Generative AI for Influence Operations** 15
- **State-Sponsored Threat Actors Are Likely Seeking to Expand Use of Large Language Models (LLMs), but Barriers to Automated Attacks Remain** 16
  - *Looking Ahead: The Future of AI Operations* 17

Tactics and Techniques Emphasize Defense Evasion 18
- **Remote Management Tools Enable Hiding in Plain Sight** 18
- **MacOS and Linux Malware Continue to Diversify** 19
  - *MacOS-Focused Infostealers and Trojans Increase in Number and Sophistication* 20
  - *Linux Systems Targeted Through Poisoned Utilities, Hypervisors, and Cross-Platform Functionality* 20
- **TTPs Involving Defense Evasion Show Greatest Increase** 21
  - *Looking Ahead: Tracking Adversarial Actions Off the Disk* 22

Reflections on 2023 Predictions 23
- **2023 Predictions** 23
- **2024 Outcome** 23

Outlook: 2025 Predictions 27

Growing Adoption of SaaS Leads to Identity Exploits

The increasing adoption of interconnected software-as-a-service (SaaS) applications in 2024 meant that compromised credentials were even more useful to cyber threat actors, who exploited gaps or misconfigurations in single sign-on (SSO) or multifactor authentication (MFA) policies to gain access to corporate ecosystems.

SaaS Applications Provide New Opportunities to Exploit Stolen Credentials

The average company uses approximately 371 software-as-a-service (SaaS) applications, which is up 39.4% from 2021 (266) and is expected to grow throughout 2025. Each of these applications typically comes with its own browser-based log-in portals and can be integrated into an enterprise single sign-on (SSO) solution. However, that integration is not always as seamless as intended if users access applications outside of the approved identity access management (IAM) environment. This rapidly growing identity attack surface almost certainly gives threat actors more opportunities to abuse stolen or exposed credentials, which are used in 77% of web application attacks. Criminals took advantage of this combination of increased SaaS reliance and exposed credentials to carry out two of the most notable attacks in 2024, namely the Change Healthcare and Snowflake breaches.

The ALPHV and RansomHub ransomware attacks affecting Change Healthcare involved the use of valid credentials to access a Citrix application that enabled remote access to desktops before pivoting to other resources. The Citrix Gateway that served as the SSO interface for remote access to Change Healthcare's network did not have multi-factor authentication enabled, allowing for the worst confluence of factors: a very common identity threat (credential abuse) being able to take advantage of a very common identity and access management measure (SSO). Once access was established, the threat actor moved laterally through the network, collected and exfiltrated patient data, and deployed the ransomware payload. Change Healthcare paid ALPHV an initial $22 million ransom, and there is evidence, but no confirmation, that it later paid a RansomHub affiliate an additional ransom following re-extortion.

The simultaneous breach of multiple Snowflake cloud storage accounts involved stolen credentials for hundreds of organizations. The financially motivated threat actor UNC5537 compromised data from multiple companies' Snowflake instances using credentials stolen via infostealer infections. UNC5537 gained access to environments not protected by MFA, similar to Change Healthcare's Citrix portal. Once they obtained access, UNC5537 used simple Structured Query Language (SQL) queries to conduct reconnaissance, identify data of interest, and exfiltrate it from the Snowflake environment. UNC5537 then sent individual extortion messages to affected organizations while selling the data on known underground and dark web forums. By carrying out these attacks simultaneously, UNC5537 created initial confusion around the source of the breach, with early reports suggesting that the Snowflake platform itself had been compromised.

Infostealer Infections Increasingly Target Personal Devices and Obtain More Credentials per Infection, Increasing the Risk to SaaS Ecosystems

In both the Change and the Snowflake data breaches, the initial valid credentials were obtained through infostealer infections. Recorded Future observed that most infostealer infections affected personal or small-to-medium business-owned devices in 2024. Personal devices are not usually subject to the same monitoring, usage restrictions, or security resources as enterprise devices, which makes it more likely that an infostealer will not be interrupted. In addition, credentials can remain exposed long after the initial infostealer infection, as demonstrated by Mandiant's finding that some of the credentials used in the Snowflake breach were stolen in 2020.

The number of credentials stolen per device has also steadily increased since 2021, likely due to the growing number of applications users routinely log in to, both in their corporate environments and personal devices. This increases the likelihood that at least one of those credentials provides access to a corporate resource. The growing number of log-ins across devices can be difficult for security teams to manage, especially if users are logging in outside the enterprise's centralized identity management system.

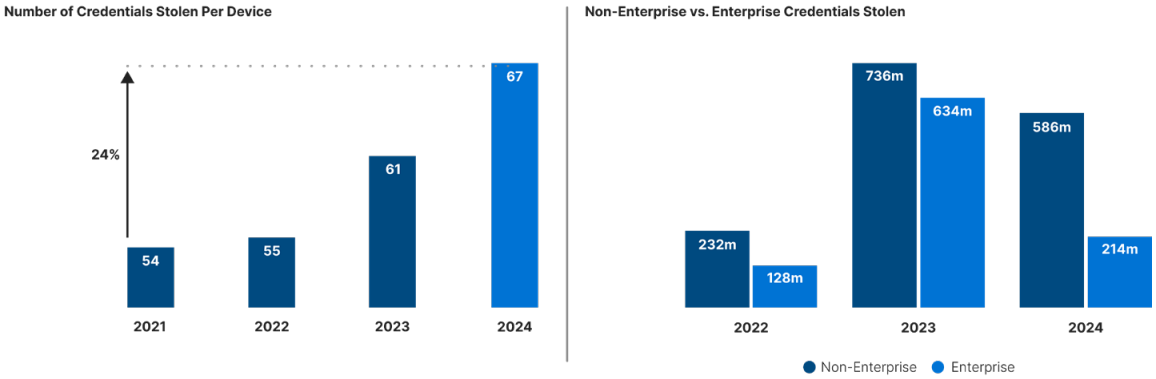

Figure 1: *Personal devices are targeted more often than enterprise devices, while the number of credentials stolen per device has increased over the last three years (Source: Recorded Future)*

Geographic Dispersion of Workforces Increases Insider Threat Risks

The adoption of SaaS and virtualized cloud environments has been in part driven by the need to connect on-premise and remote working environments. However, remote access has also enabled threat actors to take advantage of not just stolen credentials but entire stolen identities to gain access to sensitive company resources. In December 2024, the Federal Bureau of Investigation (FBI) indicted fourteen North Korean nationals for allegedly using stolen identities to pose as remote workers to obtain jobs at United States (US) companies. One cybersecurity company that unknowingly hired one of these workers reported that its new hire used a "valid but stolen US-based identity", allowing him to pass background checks, four video-based interviews, and other pre-hire vetting.

Looking Ahead: Taking Steps to Secure Cloud Identities

An estimated 22% of jobs are expected to be fully remote by 2025, meaning identity management from hiring through termination is likely to remain a complex endeavor. As Insikt Group noted last year, hybrid workspaces are increasingly moving more companies to depend on third-party infrastructure and software to support business operations. Last year, Insikt Group observed the threat of mass vulnerability exploits — brought to prominence by the exploitation of the MOVEit file transfer vulnerability (CVE-2023-34362) — while this year featured threat actors taking a different approach to exploit organizational complexity. Multifactor authentication would have almost certainly helped prevent both the Change and Snowflake breaches, but consistent configuration across a complex and often cloud-based infrastructure is easier said than done.

In response to the growing threats to SaaS applications, companies are already increasingly investing in what they hope are more mature or less easily targeted identity management solutions. One study found a 400% increase in passkey adoption in 2024, with over 100 popular apps now offering passkey support. These investments appear to be paying off, as 83% of companies reported that their identity security investments have helped reduce identity-related risk. However, proper configuration and consistent maintenance are essential to ensuring these solutions work effectively, as criminals will continue to look for new ways to exploit gaps in protection.

Extortion Groups Proliferate Despite Law Enforcement Action

While law enforcement actions in 2024 temporarily disrupted the operations of LockBit and major infostealers, phishing kits, and DDoS services, criminals adapted by reorganizing into smaller groups outside of major RaaS affiliates, demonstrating operational resilience.

Law Enforcement Actions Disrupt Two Major Affiliates

The ransomware landscape saw two major shifts in 2024 due to the disruption of LockBit and the departure of ALPHV, two major ransomware-as-a-service (RaaS) groups. In February 2024, law enforcement agencies in the US, United Kingdom (UK), and several European countries announced the disruption of LockBit's operation in a concerted operation dubbed "Cronos", seizing several of the group's network resources, arresting two affiliates, and freezing cryptocurrency accounts. Further operations continued in October 2024, when law enforcement revealed additional actions undertaken against LockBit's operations. These actions included arresting five additional individuals linked to LockBit and seizing nine servers that were part of the group's infrastructure. Meanwhile, ALPHV (also known as BlackCat) took its affiliate network offline in March 2024, shortly after collecting a $22 million payment from Change Healthcare. Researchers suspect the sudden shutdown was part of an exit scam to avoid paying affiliates amid increasing law enforcement attention.

The disruption of LockBit and ALPHV's activities was significant. According to Recorded Future data, Lockbit accounted for 23% of all ransomware activity in 2023, making it one of the most active groups in the previous year. ALPHV also appeared among the top five most active ransomware groups last year.

Law enforcement actions did not only focus on RaaS in 2024. Other notable takedowns included "Operation Endgame", which shut down infrastructure supporting at least four malware groups, including IcedID, Smokeloader, Pikabot, and Bumblebee. US law enforcement also issued indictments for members of the criminal group Scattered Spider and the hacktivist group Anonymous Sudan, the latter of which also sold access to its distributed denial-of-service (DDoS) tool called the Distributed Cloud Attack Tool (DCAT) in addition to conducting attacks directly.

Despite law enforcement actions on the ransomware and malware-as-a-service ecosystem, overall volumes for reported attacks have not decreased significantly. Ransomware groups adapted to the disruptions by reorganizing and rebranding. Recorded Future data shows that ransomware victims posted on LockBit's extortion website decreased steadily over the last year, dropping from 205 posts in Q1 to just four posts in Q4. However, the LockBit malware strain continued to be used by various emerging groups, likely due to the ongoing use of the LockBit 3.0 ransomware builder, which was leaked in 2022. Examples of groups using LockBit malware include Dragon Force, a hacktivist threat actor turned ransomware group, and CosmicBeetle, a minor cybercriminal threat actor Recorded Future has tracked since 2020.

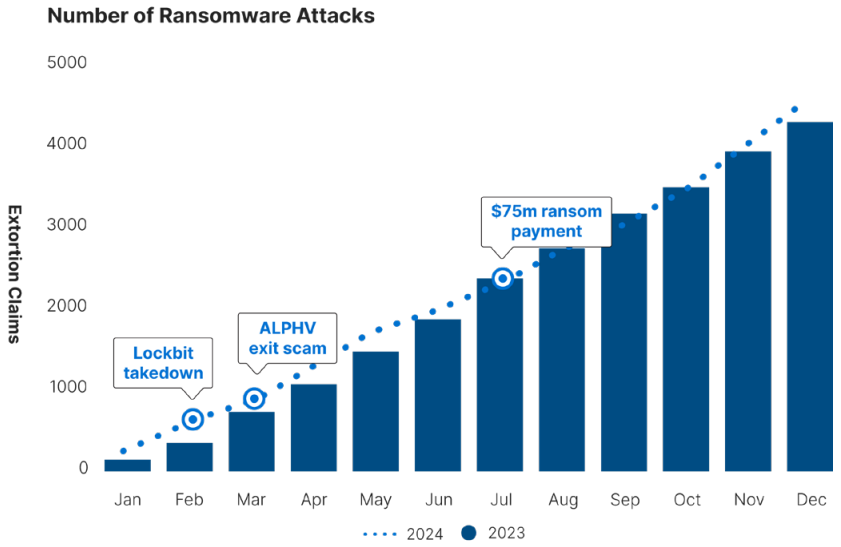

Figure 2: *While the velocity of attacks decreased slightly in August 2024, between September and October 2024, attacks resumed the pace set over the last two years (Source: Recorded Future)*

Additionally, the number of new groups and extortion blogs has increased throughout 2024. In 2023, Insikt tracked 32 new ransomware groups for the entire year. Meanwhile, in the summer of 2024 (June to August), there were fourteen new blog sites and 62 new variants, with the most prevalent group (RansomHub) only representing 12% of total extortion posts. In the previous six months, by comparison, LockBit alone represented 35% of all extortion posts. This activity indicates a pivot from consolidated, high-profile ransomware-as-a-service models toward a more fragmented criminal ecosystem where any one group is less likely to attract law enforcement attention. The increased number of observed new ransomware variants is likely facilitated by leaked source code and builders becoming more widely used and shared among threat actors.

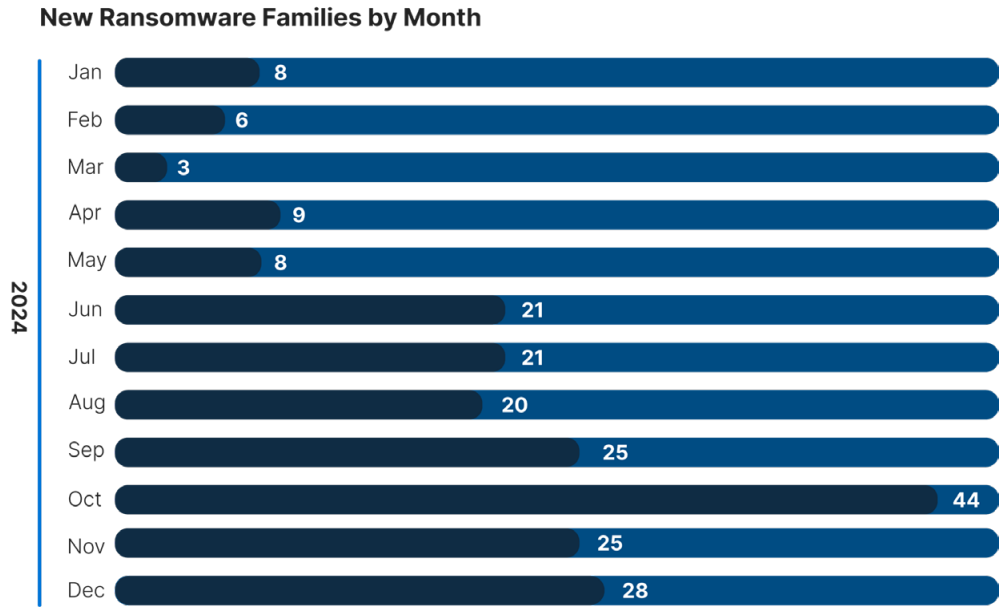

Figure 3: *The number of new ransomware families detected each month has significantly increased throughout H2 2024 (Source: Recorded Future)*

Ransomware persists because it remains a lucrative criminal endeavor, with cumulative ransom payments reaching $459.8 million through June 2024, setting 2024 on track to be the highest-grossing year yet for ransomware payments, according to Chainalysis. Chainalysis observed that in 2024, ransomware groups extorted a smaller number of significantly larger ransom payments against high-profile organizations, a practice known as "big game hunting". The year 2024, for example, saw the largest ransomware payment ever recorded: approximately $75 million extorted by Dark Angels from a Fortune 500 company. The previous record was a $40 million payment from CNA Insurance in 2021.

In addition to remaining profitable, ransomware operators often see few consequences despite Western law enforcement action. Disruptions tend to have a short-term impact on the ransomware economy, with ransomware victim extortion posts on the dark web recovering quickly, as seen in the chart below. Indictments have even less impact if the target lives and works in a country from which they cannot be extradited to the US or Western Europe — such as in one of the countries making up the Commonwealth of Independent States (CIS). While the Russian government does not often punish cybercriminals who target outside their areas of influence, there have been at least two cases this year where arrests were made involving threat actors who had launched attacks against US entities.

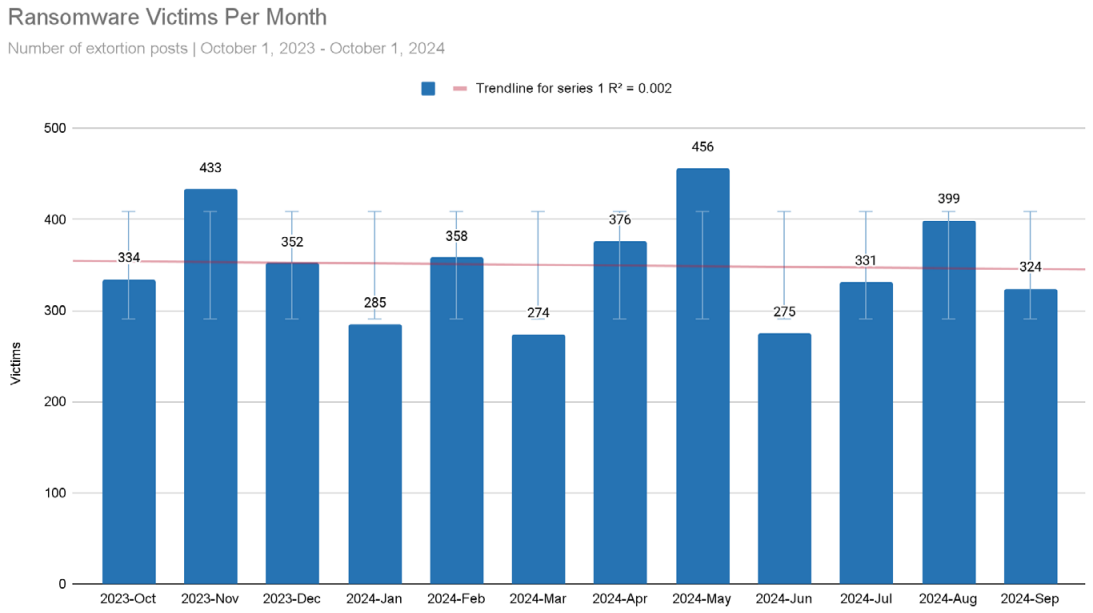

Figure 4: *Despite law enforcement actions throughout the year, there has been limited effect on the number of ransomware victims posted on extortion blogs each month (Source: Recorded Future)*

Looking Ahead: Adapting to Variety in the Ransomware Ecosystem

As certain groups rebrand or rebuild using leaked source code or by acquiring code from other ransomware groups, they can adapt and expand their arsenal. Inter-group collaborations, rebranded operations, and the rise of smaller, independent threat actors fill the void left by centralized groups, leading to an expanded variant ecosystem. A larger pool of variants presents emerging challenges for defenders as proactive threat hunting and attribution become more difficult. Defenders are less likely to find success with YARA rules that focus on distinct syntax within the code base, as they are more likely to vary between attacks. However, the talent pool for building distinct functionality or attack chains is unlikely to increase as rapidly, meaning heuristic detection of common precursors to ransomware encryptor execution will remain consistent.

Industry Analysis Gives Insight into Threat Actor Priorities

Despite the changes in cybercriminal operations described above, sector-specific targeting patterns have remained consistent over the last few years for both extortion attacks and stolen databases. Manufacturing, healthcare, and construction were the most targeted industries by ransomware groups based on Recorded Future's ransomware victimology data, with the manufacturing industry leading for the third year in a row.

Threat Actors Seek to Amplify Urgency to Increase Extortion Payouts

For manufacturing, a combination of limited IT and security investments, the high cost of operational downtime, and the complexity of IT and operational technology (OT) networks are likely to result in ransomware operators perceiving greater returns for attacks in this industry. Furthermore, the manufacturing industry is one of the largest industries globally, presenting the most targets of interest to threat actors.

- **Inconsistent network segmentation allows disruptions in operational technology:** Manufacturing companies tend to have complex environments using IT and OT systems that are not always protected by similarly dense security measures. For example, a 2022 study of British chemical companies demonstrated that 74% of companies reported that their OT environment was accessible from corporate IT networks. A year later, OT security company Dragos reported that 70% of attacks affecting OT originated in IT systems.
- **Operational disruption leads to high costs of downtime:** The cost of downtime for manufacturers reportedly ranges from $8,662 to $33,333 per minute, with large automotive companies suffering costs on the higher end. The average downtime following a cyberattack varies wildly, from several hours to over four months.

The healthcare industry is also characterized by a low tolerance for downtime due to the potential for adverse effects on patient health and safety. In May 2024, Ascension, a private healthcare provider in the US, suffered a ransomware attack that disrupted operations across many of its 136 hospitals, forcing the diversion of ambulances and closure of pharmacies, with critical IT systems offline for six weeks. In addition to monetary costs, ransomware attacks on hospitals have been shown to increase emergency room wait times — not just at the affected hospital but also at other hospitals in the region. The direct effects on patient care and, potentially, patient lives add further urgency to resolving ransomware incidents at healthcare facilities, putting additional pressure on companies to pay up.

The types of real-world effects of shutting down a factory or hospital floor are not the only factors driving urgency in ransomware payouts. Targeting a key third-party vendor or service provider can result in knock-on service disruption across an industry. In June 2024, CDK Global shut down data centers, phones, and applications following a BlackSuit ransomware intrusion. In turn, these mitigation measures disrupted services for approximately 15,000 car dealerships across North America. It has been widely reported that CDK Global paid a $25 million ransom. US-based economic consultant Anderson Economic Group estimated that CDK Global's shutdown cost auto dealers over $600 million in just two weeks. According to the Detroit Free Press in August 2024, German manufacturer AEG estimated that car dealers' total direct losses over the three weeks of the cyberattack reached $1.02 billion, with an over 5% drop in car sales for June 2024 when compared to the previous year. The high costs and widespread disruptions added urgency to the threat actor's extortion demands, very likely factoring into CDK Global's decision to pay.

Value of Stolen Data Driven by Future Monetization

In addition to extorting companies with the threat of data exposure, criminals monetize access to victim networks through selling stolen data or direct access to networks. According to Recorded Future data, the number of posts from initial access brokers (IABs) on dark web criminal forums in 2024 is roughly equal to the number from 2023. However, breaking out posts by type of service offered shows that the number of exposed databases is increasing, while direct access to compromised organizations appears to be on a downward trend from a peak in 2022. The most represented industries in these criminal advertisements were consumer goods, representing 13% of all posts, followed by technology and government/public sector (both at 7% of all posts).

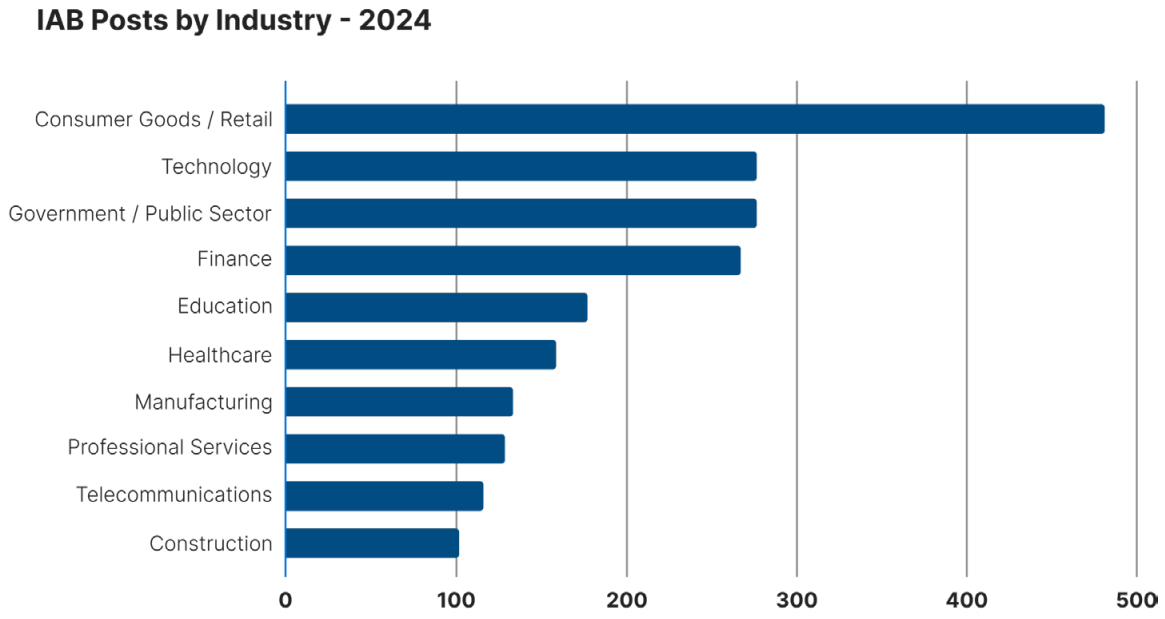

Figure 5: *Consumer goods/retail organizations are the most represented data for sale on criminal forums (Source: Recorded Future)*

A closer look at the prices of advertised databases shows that less frequently referenced industries tend to command higher prices. For example, databases from the retail industry were the most common, and their median price was $320 per listing, below the overall median price of $500. (Insikt Group's use of median pricing is due to outlier numbers in the dataset.) The median price of healthcare databases and telecommunications databases increased significantly between 2023 and 2024 (100% and 163% increases, respectively); telecommunications featured the most expensive industry data with a median price of $1,000 per listing while representing only 3% of total listings. Databases stolen from the education sector saw one of the most significant price increases, jumping from a median price of $298 in 2022 and then $43 in 2023 to $700 in 2024.

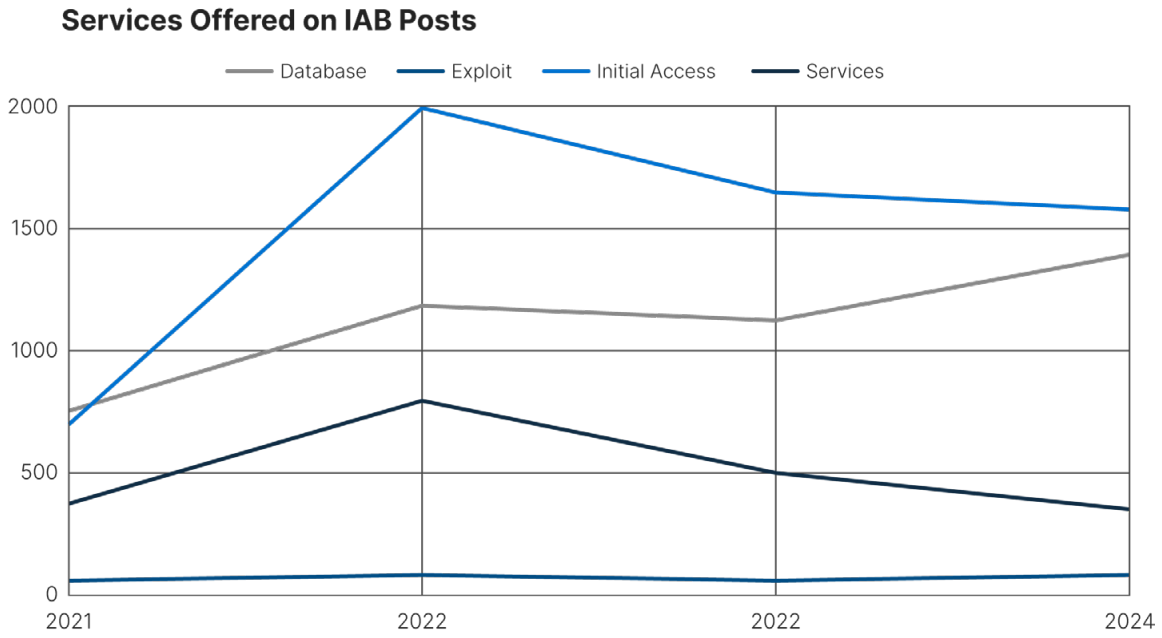

Figure 6: *Despite a peak in 2022, initial access offerings represent only slightly more posts than databases on criminal forums (Source: Recorded Future)*

Prices of databases are driven by multiple factors, including the quality of the data, the profile of the victim organization, or the potential to monetize the data further for identity theft and other scams. Broader reports about the identity theft landscape can help contextualize the utility of exposed data. According to the US Federal Trade Commission, credit card fraud was the most common type of identity theft in the first three quarters of 2024. While financial sector data is highly likely to contain useful information for credit card fraud, other sectors, including healthcare, utilities, and educational institutions, collect personal, immutable data such as previous addresses, family members' names, or Social Security numbers. Telecommunications customer data is also likely useful for carrying out SIM swap fraud, which continues to be a pervasive method for bypassing two-factor authentication (2FA) and committing other types of fraud.

Looking Ahead: Anticipating the Highest "Return on Infection"

Understanding how criminals monetize access to networks, assets, and data can help inform mitigation strategies. While financially motivated threat actors do not target specific companies in the same way state-sponsored threat actors do, target prioritization is likely to remain focused on industries with the highest "return-on-infection", which in 2024 meant the manufacturing, healthcare, and financial industries by dollar amount. This is roughly consistent with Recorded Future's observations, where the most ransomware extortion posts involved companies in the manufacturing, healthcare, industrial equipment, construction, and service industries.

Escalating Global Hostilities Drive Critical Infrastructure Targeting

Threat actor groups associated with Iran, China, and Russia, as well as their hacktivist proxies, targeted civilian critical infrastructure in disruptive and destructive cyberattacks, taking advantage of deniability to advance their objectives through hybrid conflict.

Chinese State-Sponsored Pre-Positioning Demonstrates Capabilities to Disrupt US Critical Infrastructure

In February of 2024, senior US cybersecurity officials warned that the Chinese state-sponsored threat actor known as Volt Typhoon was seeking to pre-position itself in US critical infrastructure networks. Per the warnings, "pre-positioning" intends to ensure threat actors have access to critical networks to carry out a strategically timed cyberattack in the event of escalating geopolitical conflict. Reports of Volt Typhoon's activity date back to 2023, when Microsoft detected activity at a US organization with the likely intent of "pursuing development of capabilities that could disrupt critical communications infrastructure". The observed Volt Typhoon activity on critical infrastructure networks represents a shift in Chinese cyber threat activity against the US, which has traditionally been characterized by industrial or geopolitical espionage. Other examples of pre-positioning activity likely occurred in the context of the conflict on the China-India border, where Indian power plants continue to be targeted by reconnaissance activity despite a recent de-escalation of tensions.

Successful pre-positioning activity, as well as the continued development of advanced espionage techniques (see "Salt Typhoon" below), are the result of a nearly decade-long effort to "shift to stealth." The Chinese government has invested substantially in developing stealth techniques, including researching zero-day vulnerabilities and exploiting edge devices, the latter of which was detected in an espionage campaign primarily targeting Taiwanese entities.

The pre-positioning activity throughout 2024 occurred in the context of a years-long trend toward increasingly adversarial relations between the US and China. The US and China continue to clash on China's territorial expansion in the South China Sea and the Chinese government's territorial claims over Taiwan, as well as over human rights, technology transfer, trade, and espionage. Further, the Trump administration has nominated several individuals to key national security and diplomatic positions who have been outspoken about countering China's growing global influence.

Salt Typhoon's Massive Telecom Hack and US-China Relations

In October 2024, the Wall Street Journal reported that Chinese hackers had breached US telecommunications companies, gaining access to metadata associated with personal communications and systems used in court-ordered wiretapping operations. Further details about the extent of the breach emerged in briefings to the US Congress, where it was revealed that dozens of telecommunications companies were targeted and the intrusion was still ongoing. The threat actors, dubbed "Salt Typhoon", reportedly used access to the companies to target the phones of high-profile political figures. This activity aligns with previous Chinese espionage activity targeting telecommunications companies in Asia and the Middle East to reach intelligence targets. Additionally, evidence from the February "i-Soon Leaks" indicated an interest in targeting call data records from telecommunications companies worldwide.

While espionage operations are distinctly intended to not disrupt their target networks, the discovery of such a wide-ranging infiltration of US critical infrastructure networks degraded US-China relations. As of this writing, the reaction from Congress has primarily focused on the security lapses at the telecommunications companies themselves, such as the presence of outdated routers and a lack of monitoring capabilities. Post-incident investigation of the Salt Typhoon breaches has also focused on perceived shortcomings related to federal cyber agencies' access monitoring and incident response, though the specific details remain classified. Past Congressional outrage over cyberattacks has led to new cyber legislation, such as the passage of the Cyber Incident Reporting for Critical Infrastructure Act following the Solarwinds breach. However, the Biden administration also imposed sanctions to punish Russia for the breach. Sanctions or other punitive measures toward China will also likely emerge from this activity, further straining US-China relations.

Russia-Linked Sabotage Efforts Indirectly Advance War Aims

Similarly, Russia has sought to advance its war aims in Ukraine by indirectly targeting Ukraine's allies. In early May 2024, the Cybersecurity and Infrastructure Security Agency (CISA) alerted on hacktivist activity primarily targeting vulnerabilities in the water and wastewater sector. Pro-Russia hacktivists have been observed remotely manipulating human-machine interfaces (HMIs) to cause low-level disruption or damage. CISA reports that many of these incidents were due to vulnerable internet-facing devices or the use of weak or default passwords. The Russian state-sponsored threat group Sandworm has also used hacktivist personas and proxies to conduct hack-and-leak operations and amplify the impact of destructive cyberattacks through influence operations. In December, the hacktivist group Xaknet, linked to Sandworm, claimed responsibility for disrupting Ukraine's state registers, stealing data, and preventing citizens from accessing digital services. Sandworm's control of and relationship to its hacktivist proxies likely varies depending on the group, though Mandiant assesses that the closest operational relationship is with RussianCyberArmy_Reborn.

In addition to calling out cyber disruptions, North Atlantic Treaty Organization (NATO) countries have increasingly been calling attention to Russia's "shadow war", which deploys destructive sabotage operations against European critical infrastructure. Russia's sabotage operations almost certainly seek to destabilize NATO allies, degrade NATO's war-fighting capacity, and disrupt NATO support to Ukraine (such as targeting military assets NATO committed to providing to Ukraine), among other objectives. Insikt Group identified at least three attacks in June that were plausible Russian sabotage efforts: a break-in at a water treatment facility in Finland, an explosion at an arms factory in Poland, and a fire at a railroad terminal in Poland. Insikt Group also identified 21 additional incidents that occurred the same month throughout Europe that could also have been potential sabotage events, but more information is needed to confirm.

Looking Ahead: Geopolitical and Cyber Convergence

Russia's activity through physical sabotage and "hacktivist"-driven cyberattacks represent "strategic information attacks", the use of non-kinetic means to disrupt and destabilize adversaries. To advance this strategy, indirect attacks must be disruptive enough to have their desired effects without causing significant political blowback or military retaliation.

Because indirect attacks on critical infrastructure occur in the context of geopolitical conflict, changes in relations between countries will likely escalate or restrain cyber threat activity. Given the uncertainty around the foreign policy of the new Trump administration, defenders should anticipate the continuation of cyber and physical threats characterized by varying degrees of disruption and deniability. Scenarios that could accelerate malicious activity may include Chinese efforts to expand its territory in the South China Sea or the further deterioration of Russia's relations with the West –– particularly the European Union (EU), United Kingdom (UK), and the United States (US).

Generative AI Accelerates Spread of Inauthentic Content in Historic Election Year

State-sponsored malign influence operations increasingly relied on GenAI to craft and distribute deceptive ,persuasive content to influence elections worldwide.

State-Sponsored Adversaries Continue Experimenting with Generative AI for Influence Operations

Throughout 2024, over 2 billion voters worldwide headed to the polls to choose their governments, including the US, the UK, India, Indonesia, France, and the EU. The elections also provided ample opportunities for state-aligned influence actors to attempt to mold public opinion, using generative AI tools to accelerate operations. Russia, China, and Iran each engaged in malign influence operations that aimed to destabilize Western democracies, inflame political and social tensions, and sow discord among populations to advance their own strategic geopolitical goals. GenAI increases the speed and volume at which adversaries can launch and execute influence operations on social media networks. Despite the improved efficiency, campaigns generally fell short of making content believable to the target audience. As seen with Operation Overload, however, believability is not always the primary objective — overwhelming the information environment with inauthentic or junk content to degrade the overall quality of discourse can be an end in itself.

Russia, in particular, has evolved its use of generative AI to produce more human-like and persuasive content and enhance the scale and scope of influence operations supporting broader state-run campaigns. For example, Russian-operated "Operation Undercut" used at least 500 social media accounts to spread AI-enhanced videos promoting various Russian narratives that targeted the EU and EU citizens. According to the US Department of Justice (DOJ), Russian state-aligned media organization RT used a covert tool designed to create and control an AI-enhanced bot farm called Meliorator. The tool uses GenAI to create more realistic networks of artificial personas with account attributes more aligned with legitimate accounts than a bot account. Meliorator was likely intended to act as a force multiplier, enabling the creation and management of multiple realistic-looking accounts in bulk.

Some networks aimed to sway electoral outcomes, undermine public confidence, and sow global discord. The Russia-aligned influence operation "Operation Overload" conducted a malign influence campaign designed to sow doubt, confusion, and discord during the 2024 US presidential election cycle, the July 2024 French elections, and the 2024 Paris Olympic Games. A key feature of this campaign was its velocity and volume of content, including AI-generated voiceovers impersonating authentic news organizations aimed at overwhelming the finite investigative resources of media organizations, researchers, and civilians who encountered the information. By creating an overabundance of inauthentic persuasive content and then automating its distribution through social media, Operation Overload forces defenders to spend more time collecting evidence and verifying the inauthenticity of the information instead of verifying legitimate, time-sensitive leads.

Iran, meanwhile, used an influence network Insikt Group refers to as "Emerald Divide", which aimed to foment anti-government sentiment among the Israeli population. Emerald Divide demonstrated the ability to pivot in response to the changing political environment in Israel, tailoring AI-generated content to the conflict between Israel, Hezbollah, Hamas, and Iran. Emerald Divide used a social media network that included at least seven primary accounts and a network of over 250 accounts engaging in coordinated inauthentic behavior (CIB). This included a campaign where inauthentic accounts purporting to be Israeli rabbis published deepfakes containing anti-LGBTQ+ content, while other Emerald Divide-controlled accounts criticized the same content. The accounts stoked division in both pro-Israel and pro-Hamas circles and were distributed through the CIB network. Insikt Group assesses that Emerald Divide likely succeeded in influencing its target audience based on observable results, including anti-government protests and posters observed in the target area of interest.

Lastly, Chinese influence actors promoted AI-generated content that discussed controversial US issues and criticized the Biden administration. Chinese influence operations attempt to amplify division overall and portray democracies as a chaotic and less desirable political system. Their social media personas often highlight instances of perceived incompetence or misalignment between administration goals and public needs. Beyond the US election, China amplified favorable narratives around specific events or issues that otherwise would appear unfavorable, such as redirecting hashtags associated with the Tiananmen Square massacre to content highlighting police violence in the US. Chinese influence operations generate content in over 40 languages, demonstrating the international aims of its sponsors.

State-Sponsored Threat Actors Are Likely Seeking to Expand Use of Large Language Models (LLMs), but Barriers to Automated Attacks Remain

In addition to inauthentic content, threat actors associated with China, Russia, Iran, and North Korea have attempted to expand the use of AI into other malicious applications throughout 2024. OpenAI confirmed that state-sponsored threat actors used the platform for malware development, social engineering reconnaissance, influence operations, and spearphishing lure generation. Other evidence that cyber threat actors are using GenAI in malware development has already emerged through the discovery of visual basic script (VBScript) and JavaScript code containing artifacts indicating the use of the technology.

However, GenAI has not been successfully used to develop or deliver fully automated malware. Security researchers have developed proof-of-concept (PoC) malware through AI prompt engineering and training. Insikt Group created a variant of the PowerShell-based infostealer STEELHOOK that used an LLM to reason through its environment, making in-memory code updates to bypass YARA rules. Similarly, Hyas released two PoC pieces of malware, EyeSpy and BlackMamba, which used LLMs to dynamically generate polymorphic behavior. The LLM synthesizes the malicious code that drives the keylogger at runtime, evading signature-based detections.

All this being said, the countries above are actively investing in AI capabilities for more automated offensive operations. China, in particular, is expanding its military, private, and academic AI capabilities in pursuit of enhancing offensive operations in the physical and cyber domains. Russia, Iran, and North Korea are primarily in the nascent or experimental stages of AI integration, focusing more on exploring the potential of AI within their cyber strategies but lacking corresponding investment in the private sector entities that drive the majority of innovation in the area.

Looking Ahead: The Future of AI Operations

It remains likely that adversarial nation-states will primarily use GenAI for generating inauthentic content rather than malware deployment or other offensive capabilities in 2025. Polymorphism, mutex mechanisms, hiding malicious behavior in memory, and other behaviors these LLMs synthesize are not new, and existing mechanisms are designed to detect these behaviors. To generate a malicious script or program, LLMs need to train on existing malware and kill chains to understand the mechanics of a cyberattack. Currently, LLMs are not sophisticated enough to invent entirely novel techniques based on the available training data. Further, as with other applications of generative AI, the outputs are generally less complex than pure human outputs. AI applications will likely be used to enhance or troubleshoot threat activity, as observed by OpenAI throughout 2024.

At the same time, efforts to limit the production of politically motivated AI content will likely be deprioritized in the coming years. In the US, Insikt Group anticipates diminished public reporting on inauthentic content networks. In line with free speech concerns, the Trump administration issued a day-one executive order prohibiting the federal government from infringing free speech by labeling domestic speech as “mis-” or “disinformation”. That said, the efficacy of research into inauthentic content has always correlated to the will and ability of social media platforms and hosting companies to remove abusers, so it will be the decisions of the private rather than the public sector that will determine how much politically motivated AI content is allowed to spread over the next few years. While international and US state laws have been passed against AI-generated political content, these have had limited impact on the development and spread of inauthentic content.

Tactics and Techniques Emphasize Defense Evasion

Tactics, techniques, and procedures (TTPs), once limited to sophisticated, state-resourced threat actors, are increasingly becoming common in criminal operations, including a significant jump in observed defense evasion techniques.

Remote Management Tools Enable Hiding in Plain Sight

Threat actors' abuse of remote monitoring and management (RMM) tools surged in 2024, driven by their capability to evade detection while maintaining operational efficiency. Traditionally used for IT support, these tools have become increasingly adopted by attackers aiming to bypass endpoint detection and response (EDR) mechanisms. Huntress reported seeing a 214% increase in incidents involving RMM tools between January and October 2024. Recorded Future data corroborates this trend, with over 600 references to RMM tools and cyberattacks in 2024 versus just 50 references in 2023. The references in early 2024 relate to an exploit involving an authentication bypass flaw in ScreenConnect tracked as CVE-2024-1709 that was widely exploited after the release of a proof-of-concept (PoC). The second major spike in references in 2024 related to BlueBravo's attack on TeamViewer, which led to the exfiltration of employee data and encrypted passwords to their IT environment.

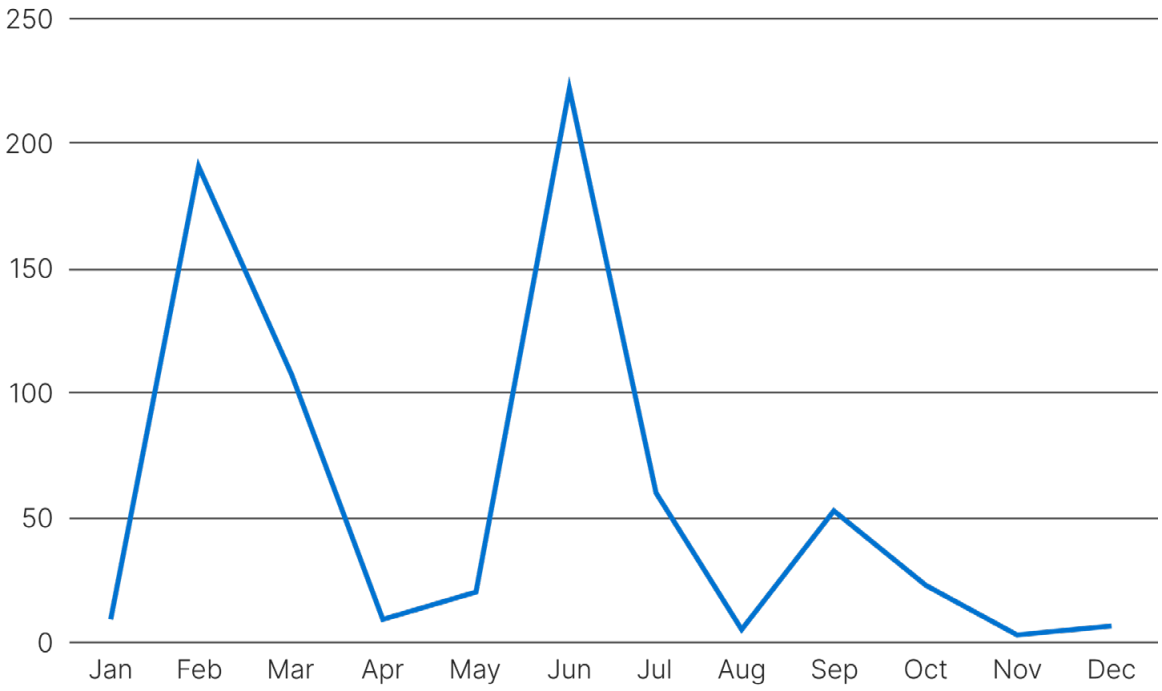

Figure 7: *Cyberattack reports mentioning remote monitoring and management tools demonstrate a significant spike in the summer of 2024 (Source: Recorded Future)*

Threat actors have delivered RMM through social engineering, such as BlackBasta's campaign to convince targets to download AnyDesk or TeamViewer in an IT help desk scam. Once the RMM program is executed on the system, threat actors can abuse that access to establish persistence and deliver other payloads.

The increased use of RMM is likely driven by two trends: increased adoption of endpoint detection and response (EDR) tools and virtualization of hybrid and remote work environments, which depend on RMM tools such as ConnectWise, AnyDesk, or TeamViewer to enable remote IT. Improvements in EDR capabilities have driven threat actors to adopt tactics that mask their activity from EDR solutions, such as by abusing legitimate programs. Increased virtualization means that RMM tools are a more common part of the IT environment, making it more difficult for defenders to differentiate between legitimate and malicious activity. In addition to RMM tools, threat actors have also used trusted sites such as GitHub primarily as a means of malware payload delivery, as GitHub pushes tend to be part of typical workflows. Other ransomware operators simply disguise malware as a legitimate RMM service, like AnyDesk or ScreenConnect, to trick users into downloading the program.

MacOS and Linux Malware Continue to Diversify

While Microsoft Windows remains the predominant target for threat actors, 2024 marked a continued rise in attacks against macOS and Linux systems, with a significant expansion in multiple types of malware. This trend aligns with the observed growth in Apple device adoption in enterprise environments and many network-related equipment or systems operating on Linux hypervisors. Intel471 identified over 40 threat actors targeting macOS devices on dark web forums between January 2023 and July 2024. From January 2024 through July 2024, 21 threat actors expressed interest in macOS-specific malware or services, which is the same number observed throughout all of 2023. This finding is echoed by Group-IB, which observed a five-fold increase in the underground sale of infostealers targeting macOS users. Moreover, threat actors are increasingly developing tools using cross-platform languages like Rust or Go, further contributing to these observed increases. Linux systems also saw increased ransomware activity.

References to Cyberattacks Targeting Operating Systems

October 1, 2023 — October 1, 2024

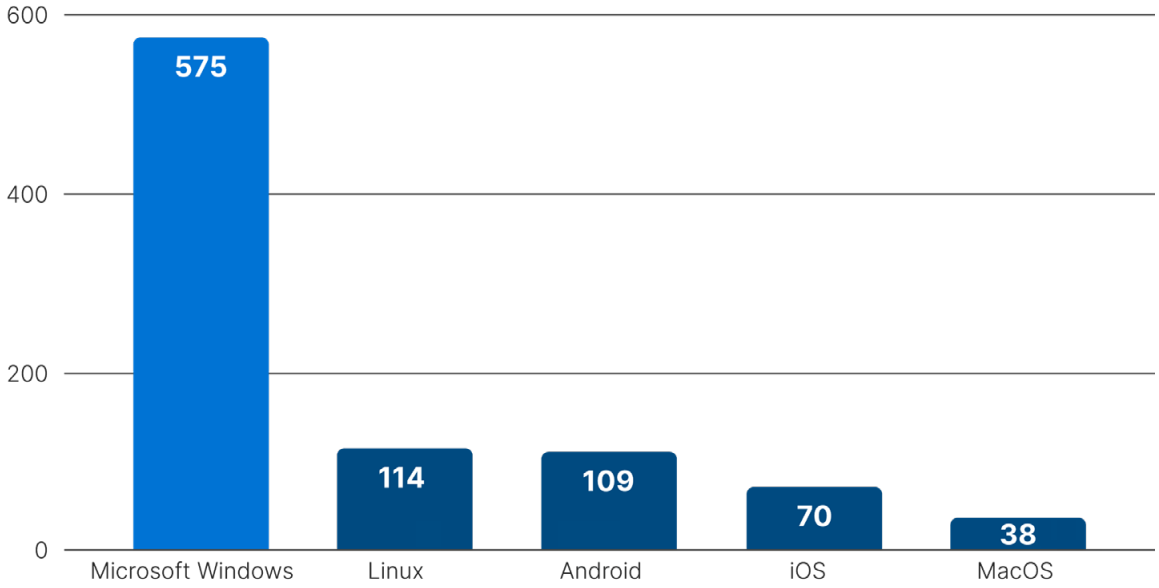

Figure 8: *While Windows remains the most commonly exploited operating system, Linux, Apple, and mobile devices are increasingly targeted (Source: Recorded Future)*

MacOS-Focused Infostealers and Trojans Increase in Number and Sophistication

Infostealers targeting macOS, such as BANSHEE Stealer, Cthulhu Stealer, and Atomic Stealer (AMOS), are becoming more common. This increase represents a perceived change in threat actor interest in exploiting macOS environments as networks or users of interest more commonly use Mac computers. HZ Rat was a macOS adaptation of a previously Windows-focused remote access trojan (RAT), as was a macOS-focused variant of the LightSpy mobile framework reported in April 2024. This variant is more refined than the iOS version discovered in 2020, as it obfuscates its command-and-control (C2) framework, among other improvements. macOS-focused trojans and other backdoors also emerged, including the Trojan.MAC.RustDoor, which was identified in February 2024. This backdoor continues the trend of threat actors writing malicious code in Rust as a way to efficiently compile into native binaries for various operating systems, including Windows, Linux, and macOS.

Cybercriminals deliver and operate macOS infostealers alongside Windows infostealers through multiple vectors, including social engineering and exploiting legitimate internet services. Criminals impersonate popular programs, such as Web3 gaming or video conferencing software, to trick users into downloading infostealers that function on multiple operating systems. Similarly, criminals have been observed using GitHub to support command-and-control infrastructure for a variety of infostealers, including AMOS. Using GitHub and other legitimate internet services enables criminals to evade detections by operating on programs that likely already exist in the target environment.

Beyond RustDoor, ZuRu emerged as a stealthy, espionage-focused macOS backdoor. Its modular architecture allowed it to dynamically load additional components based on the target's environment. These payloads enabled screen capture, keylogging, and remote shell access and used code signing and legitimate Apple developer certificates to bypass macOS Gatekeeper and XProtect defenses. The infostealer JaskaGO, similar to RustDoor, is written in a cross-platform language (Go), which allows attackers to target multiple operating systems using a single codebase, reducing development time and resources. The multifunctional infostealer and backdoor malware known as Cuckoo also emerged this year, and, like ZuRu, it is notable for its stealthy operation and modular architecture.

Linux Systems Targeted Through Poisoned Utilities, Hypervisors, and Cross-Platform Functionality

In 2024, Insikt Group observed a novel worm based on the Mirai botnet targeting Linux systems. A customized version of the malware targeted Linux-based servers, routers, web cameras, and other Internet-of-Things devices, but it functioned as a malicious cryptominer rather than a DDoS platform. In another example where malware expanded its ability to target Linux, researchers discovered BiBi-Linux, a wiper malware capable of completely erasing data on infected Linux systems. BiBi-Linux specifically targeted Israeli companies at the onset of the Israel-Hamas conflict. There is also a Windows variant of the same malware, again demonstrating that threat actors are increasingly working across platforms.

One of the products Insikt Group observed ransomware groups target the most in 2024 was VMWare's ESXi hypervisor, which is built on a Linux framework. Insikt Group observed specific variants of common ransomware payloads, including Hunters International, Play, INC, and RansomHub, that directly target the hypervisor and other Linux devices. ESXi's functionality within a network is so critical that it has remained a high-priority target. Like the macOS malware discussed above, many malware variants targeting ESXi are written in either Rust or Golang. As low-level languages, they are optimized for low-latency communication and

resource-efficient execution and can be used for process injection or hooking. Further, Rust's emphasis on memory safety reduces the likelihood of bugs, producing more reliable and stable payloads that are harder to detect or disrupt, and its memory access patterns and structures are more difficult to detect with heuristic-based detections.

TTPs Involving Defense Evasion Show Greatest Increase

The marked increases across defense evasion (TA0005) techniques — Debugger Evasion, Reflective Code Loading, and Execution Through API (application programming interface) — highlight a trend toward defense evasion strategies that do not involve writing code to the disk. Network Time Protocol (NTP)-based time evasion and reflective code loading demonstrate an adversarial emphasis on bypassing behavioral analysis and forensic detection.

Some of the most significant increases Insikt Group observed in MITRE ATT&CK techniques between 2023 and 2024 were in T1497.003 (Time Based Evasion), T1622 (Debugger Evasion), and techniques associated with data gathering that inform the evasive components of a malicious loader saw. T1057 (Process Discovery) and T1016 (System Network Configuration Discovery) are techniques Insikt Group observed in multiple attacks or malware samples where either a dropper, loader, or other binary will collect data from its environment that is then used to determine whether it is running in a debugger or sandbox environment. Certain network configurations, running processes, or other artifacts, such as the number of documents accessed in the last 100 days, can be used to discriminate between a target device and a sandbox environment.

MITRE ATT&CK TTPs with the Greatest Increase in Insikt Group® Observations

November 1, 2023 — November 1, 2024

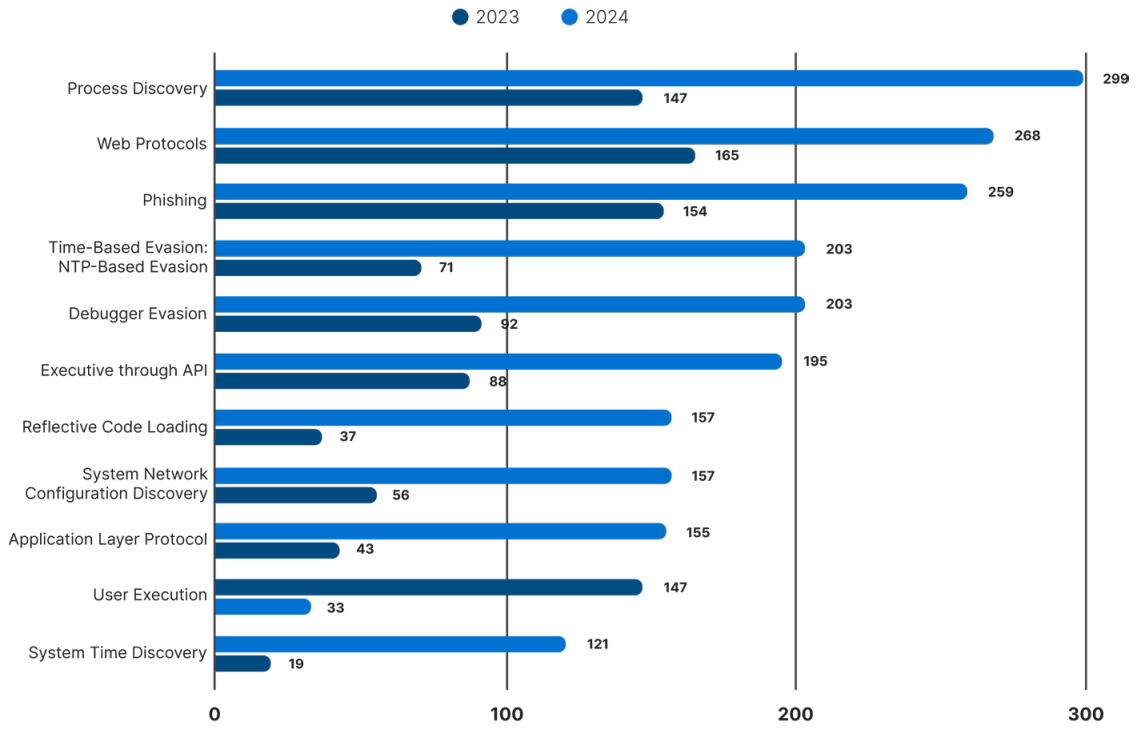

Figure 9: *TTPs with the greatest observed increase between 2023 and 2024 (Source: Recorded Future)*

In line with this report's assessment above about improved EDR use being the basis for increased abuse of RMM tools, Insikt Group also observed increases in TTPs associated with common EDR bypass techniques. T1106 (Native API) involves using API calls to execute malicious actions while blending into legitimate system activity, thereby avoiding detection by security tools. The abuse of API calls enables attackers to execute code without triggering traditional alerts and provides low-level control over system functions, helping attackers avoid reliance on detectable third-party executables. Similarly, T1620 (Reflective Code Loading) enables threat actors to load malicious code directly into memory without writing it to disk, bypassing file-based defenses common in EDR suites. Insikt Group observed that adversaries increasingly deploy memory-resident malware to evade signature-based defenses and forensic analysis, particularly among Chinese state-sponsored groups that frequently use dynamic-link library (DLL) side-loading. Tools like Cobalt Strike and custom loader scripts have also contributed to the rise in reflective code-loading use.

Looking Ahead: Tracking Adversarial Actions Off the Disk

With EDRs and other security tools becoming increasingly common, threat actors are looking for ways to move in security gaps to evade detection. These developments underline the need for defenders to adopt robust memory analysis, behavioral anomaly detection, and advanced threat intelligence integration to mitigate evolving threats. Defenders should consider more complex defenses, such as execution state mapping for known program configurations. This can interrupt binary execution or API calls, including malicious DLLs in cases where the libraries don't match approved profiles due to malicious code insertion. Execution state mapping can be combined with heuristic monitoring into a positive and negative security model that can disrupt the effectiveness of threat actor behaviors. This way, even with EDR bypass techniques like unhooking or reflective DLL loading, the malicious binary will not execute; the library would not be operating in a known state in alignment with the positive model. This can then be used to further develop the negative security model, increasing defenders' capacity to quickly detect and disrupt future threat actor behavior.

If organizations have the capacity or desire to analyze malware on their own, the growth of in-memory cyberattacks should drive how personnel are hired or trained. More generally, a greater emphasis on behavior-based detections — and especially the order of certain types of behaviors (for example, process enumeration prior to a specific process being shut down) — can contribute to identifying anomalous activity as the threat landscape continues to move away from more easily trackable indicators and binaries. This is especially true for many lateral movement techniques, such as credential dumping, a behavior that is often atypical for a normal user.

Reflections on 2023 Predictions

For our 2023 Annual Report, Insikt Group assembled a series of predictions that were believed to be either the most accurate or most worth considering based on the threat landscape at the time. However, predictions should always be re-examined after sufficient time to see how well they performed. As a result, this report provides Insikt Group's predictions from last year, along with a judgment of their accuracy.

2023 Predictions	2024 Outcome
Vulnerabilities: Given ransomware groups' (especially CL0P's) success in mass exploiting vulnerabilities in enterprise file transfer solutions to carry out thousands of successful attacks (Accellion: 2020, 2021; GoAnywhere and MOVEit: 2023), **Insikt Group predicts that at least one ransomware group will carry out a successful compromise of hundreds of targets via exploiting vulnerabilities in an enterprise third-party file transfer service** in 2024. The impact of this event will be comparable to the MOVEit campaign from 2023.	**Mostly accurate:** In the final week of the year, CL0P announced a new file transfer application exploit affecting at least 60 victim companies. While this campaign demonstrates that vulnerabilities in these applications remain particularly effective, this year's attack did not reach the same impact as last year's MOVEit exploit. Additionally, in 2024, a different threat actor claimed to have access to a tranche of data affecting at least 25 companies dating back to May 2023 — the time of the original MOVEit exploit. This incident reinforces the widespread impact of the MOVEit breach as more victims emerge.
Third-party threats: In 2024, Insikt Group expects to see software supply-chain attacks dominate the third-party threats landscape by number and severity of attacks, **with at least a 15% growth in reported incidents**. npm will highly likely continue to attract the most targeting due to its ubiquity and volume of new packets published.	**Inaccurate:** Reported cases of software supply chain (SSC) attacks remained about the same between 2023 and 2024, suggesting that while it remains a pervasive attack vector, it has not surpassed other exploits. ***Notable software supply chain incident:*** One thing Insikt Group did not expect was an elaborate social engineering attempt to insert a backdoor into a widely used, open-source software code. While code repository poisoning is not a new method of SSC attacks, the fact that the threat actor spent over two years building trust within the community in the hopes that the malicious update would go unnoticed makes this attack stand out.

2023 Predictions	2024 Outcome
Extortion groups: As more companies adopt and sustain hybrid and remote work models, **Insikt Group expects that extortion groups will increasingly target technologies supporting and securing hybrid and remote work,** especially cloud-based data storage, MFA solutions, and virtual private networks (VPNs) and that the majority of ransomware attacks will involve attacks against such assets.	**Accurate:** The increase in abuse of remote management tools, and the occurrence of multiple SaaS and cloud-based compromises, demonstrate that threat actors are increasingly taking advantage of virtualized environments to blend in with existing activity. • **165** companies affected by the Snowflake breach • **12x** increase in references to RMM tools in attacks, per Recorded Future data
Hacktivism: Insikt Group anticipates that shifts in the Russia-Ukraine war are likely to result in the **strategic diversion of hacktivist activity from groups such as Killnet and Anonymous Sudan toward the war in Gaza,** particularly as the conflict threatens to spread elsewhere in the region. Targeting of Western entities aligned with NATO and the European Union will likely continue, while focus on entities supporting Israel will likely increase.	**Mostly accurate:** Hacktivist activity did shift toward targets associated with Israel, Hamas, or Palestine, especially in the early months of the war in Gaza. However, the most significant strategic diversion Insikt Group saw in hacktivist activity in 2024 was toward making money. Insikt Group has documented Anonymous Sudan's profits from their DDoS-as-a-Service tool; in May 2024, KillMilk (formerly of Killnet) also advertised an infostealer for sale in criminal forums. One notable exception was around hacktivist activity targeting the 2024 Olympic Games in Paris to attempt to amplify attention to various political causes, though this had limited disruptive impact on the event.
Initial access methods: In 2024, Insikt Group expects that attackers are likely to **increasingly focus on stealing credentials and identities using techniques like password spraying and credential stuffing** as organizations bolster their security perimeters. Insikt Group also expects that the **"phishing" threat landscape will increasingly become the "spearphishing"** threat landscape as criminals gain more experience and resources to use GenAI to craft highly personalized campaigns that are difficult to detect.	**Partially accurate:** Insikt Group's prediction on credential stealing was spot-on, so much so that a whole section of this report was dedicated to that particular threat. While studies have demonstrated the potential for AI to improve phishing emails, Insikt Group has not yet seen evidence of its use to facilitate spearphishing at scale. • **50% increase** in the use of valid credentials between Q1 and Q3 2024, based on Recorded Future TTP data • **46%** of initial access TTPs reported in 2024 were phishing

2023 Predictions	2024 Outcome
Influence operations: As we enter a year of high-profile and numerous elections worldwide, Insikt Group anticipates that the public's awareness of deepfakes and disinformation operations will be more disruptive than the aims or activity of adversary-driven campaigns, particularly in highly polarized electorates such as the US. **Voters will likely be conditioned to write off unflattering images or press as artificially generated** (regardless of authenticity).	**Accurate:** The proliferation of GenAI content made it easier than ever for people to discount real photos and widely share known inauthentic content to advance a political point.
Technology: **Companies are almost certain to increasingly offer passwordless logins to users in 2024, such as access links to sign into websites and biometric-based authentication**. This shift will greatly reduce the value of certain leaked credentials for sale on the dark web and will force threat actors to innovate to find new ways to exploit passwordless security, such as the creation of fake access link emails. For money laundering and external, customer-facing payment fraud threats, increasing reliance on passwordless logins may drive a shift away from account takeover (ATO) tactics toward new account fraud (NAF).	**Mostly inaccurate:** The FIDO Alliance reports a **50% increase** in familiarity with passkeys in 2024, driven both by user frustration as well as the many security issues associated with traditional passwords. However, Insikt Group has not yet seen the secondary impacts of this shift on cybercrime, as evidenced by the multiple successful credential-based attacks this year.

2023 Predictions	2024 Outcome
Geopolitics: Should China's domestic economic performance worsen and become a more salient talking point on the global stage, it will likely use social surveillance and censorship to quell unrest in its own population. In its external relations, China **may pre-position disruptive cyber operations to signal its continued fortitude and deter its adversaries** like the US from taking advantage of its internal instability; however, it is unlikely China would "lash out", such as by initiating a noisy diversionary conflict or war. **Iran will continue to rely on a mixture of proxy warfare and cyber influence operations to sow unrest in the region**, including efforts to isolate Israel and oppose the US military's presence in the region. **Russia will very likely exploit perceived "war fatigue" among Western nations to influence public opinion ahead of elections** in the US and EU and will wait for election results to determine its next course of action in relation to its war in Ukraine and its relationship with the West.	**Accurate:** • **China:** Insikt Group has not seen any further prepositioning following Volt Typhoon's February exposure, though the discovery of extensive espionage activity in US telecommunications networks certainly sent a message on China's cyber capabilities and intent. • **Iran:** Iran suffered multiple setbacks in its proxy war against Israel throughout 2024. With key proxies Hamas and Hezbollah significantly weakened, Iran is ending the year in a much worse position than it started. • **Russia:** Resistance to expanding support to Ukraine was a theme in both the US and EU elections. However, it is difficult to determine how much of this was due to Russian influence operations or internal populist narratives.
Regulations: The increase in vulnerability exploits will **drive lawmakers to shift from regulating software safety to reforming software liability law.** This would make it easier for consumers to take legal action against software companies that produce insecure code; however, determining what counts as coding negligence will be a significant challenge for policymakers. In response to ongoing policies and regulations for AI, **AI companies will likely shift toward synthetic data** to train their models to avoid privacy and copyright issues, speed up development, and reduce the chances of data poisoning from threat actors.	**Somewhat accurate:** The EU updated its civil law to treat software like any other product when it comes to liability. This means individuals can sue companies and platforms for harm caused by a digital product, including in class action lawsuits. We will see how this affects software safety in 2025, when the first cases can be brought to European courts. In a hint that the US is considering a similar approach, the Biden administration took a first step by calling for more research into quantifying software security. These efforts are likely to continue into the next administration. On the other hand, AI companies continued to find sources of real data to mine with limited regulatory or public pushback throughout 2024.

Outlook: 2025 Predictions

- **AI Impersonation as the Next Big Attack Vector for SaaS Applications:** Insikt Group expects that a major breach will very likely result from one of two AI-associated factors: implementation of GenAI into enterprise workflows or abuse of AI for effective impersonation. In either case, it is likely that a SaaS application will play a role in initial access or data exposure. In the last year, companies like OpenAI (with SORA), Meta (with MovieGen), and Google (with Veo 2) have released models that make incredibly realistic fake videos and images. These tools will make it much easier for scammers to carry out existing scams in more convincing ways, such as video calls to IT help desks, in order to trick their way into gaining access to sensitive data and systems.
- **Additional Typhoon Activity Announced, Affecting New Sectors:** Despite already major revelations of intrusions into US critical infrastructure by Chinese-linked advanced persistent threat (APT) groups in 2024, Insikt Group believes that additional high-profile breaches of critical infrastructure by Chinese APTs will be disclosed in 2025. We predict that other industries beyond the energy and telecommunications sectors will disclose breaches attributed to Chinese APTs that were likely conducted with the goal of pre-positioning for disruptive operations.
- **MacOS and Mobile Threats Join Windows and Cloud:** In malware and vulnerability trends, Insikt Group predicts that one of the high-impact cyber incidents of 2025 will likely be associated with either macOS malware or mobile malware. We believe that this will result from certain environmental factors reaching breaking points, such as higher attention to macOS targets and increased access to sensitive corporate and financial data via mobile devices.
- **Crypto Fraud Will Lead to a Market-Destabilizing Event:** The soaring value of cryptocurrency, as well as a new US administration set to pursue crypto-friendly policies, will drive more aggressive and ambitious fraud attempts. Cryptocurrency scams are currently one of the most common and lucrative types of investment scams, and we predict that criminals will likely be emboldened by the crypto boom to carry out market-destabilizing scams that will at least temporarily reduce the value of cryptocurrencies and lead to calls to restrict their use.
- **Developers Embrace AI to Accelerate the Transition to New Code:** As a result of multiple factors, including improved AI coding capabilities and a new push toward software liability in Europe, Insikt Group expects that both companies and threat actors will very likely rely on generative AI increasingly to accelerate the transition to modern code libraries — whether memory-safe code for companies, or more modular code for malware.
- **The US Moves toward Cyber Regulation Harmonization:** Insikt Group anticipates that the new administration's deregulatory agenda, combined with a Supreme Court decision that puts more pressure on Congress to be clear about their intentions for how regulations are implemented, will likely drive support for passing the Cyber Regulatory Harmonization Act. Simplifying the US's complex cyber regulatory landscape has long enjoyed bipartisan support, and taking action to improve the nation's cybersecurity in the aftermath of major cyber incidents would be seen as an early win for Congress.

- **Taiwan's Critical Infrastructure Will Be at Risk:** China continues to pursue incremental escalation in its coercion of Taiwan short of initiating an armed conflict, and a disruptive cyber campaign targeting Taiwan's energy, transportation, or financial sectors is a logical next step. This would build on past small-scale disruptive actions against these sectors in 2020, late 2021, early 2022, and (arguably) late 2022, as well as activity targeting the latter two sectors in 2024. If not a disruptive campaign, Insikt Group anticipates the disclosure of widespread pre-positioning in Taiwan networks. More generally, China will continue to use its cyber capabilities to reconnoiter the policies and preparations of Taiwan and potential partners in a conflict such as the US and Japan.

Recorded Future reporting contains expressions of likelihood or probability consistent with US Intelligence Community Directive (ICD) 203: Analytic Standards (published January 2, 2015). Recorded Future reporting also uses confidence level standards employed by the US Intelligence Community to assess the quality and quantity of the source information supporting our analytic judgments.

About Insikt Group®
Recorded Future's Insikt Group, the company's threat research division, comprises analysts and security researchers with deep government, law enforcement, military, and intelligence agency experience. Their mission is to produce intelligence that reduces risk for customers, enables tangible outcomes, and prevents business disruption.

About Recorded Future®
Recorded Future is the world's largest threat intelligence company. Recorded Future's Intelligence Cloud provides end-to-end intelligence across adversaries, infrastructure, and targets. Indexing the internet across the open web, dark web, and technical sources, Recorded Future provides real-time visibility into an expanding attack surface and threat landscape, empowering customers to act with speed and confidence to reduce risk and securely drive business forward. Headquartered in Boston with offices and employees around the world, Recorded Future works with over 1,800 businesses and government organizations across more than 75 countries to provide real-time, unbiased, and actionable intelligence.