Cybersecurity Forecast 2026

Google Cloud Security

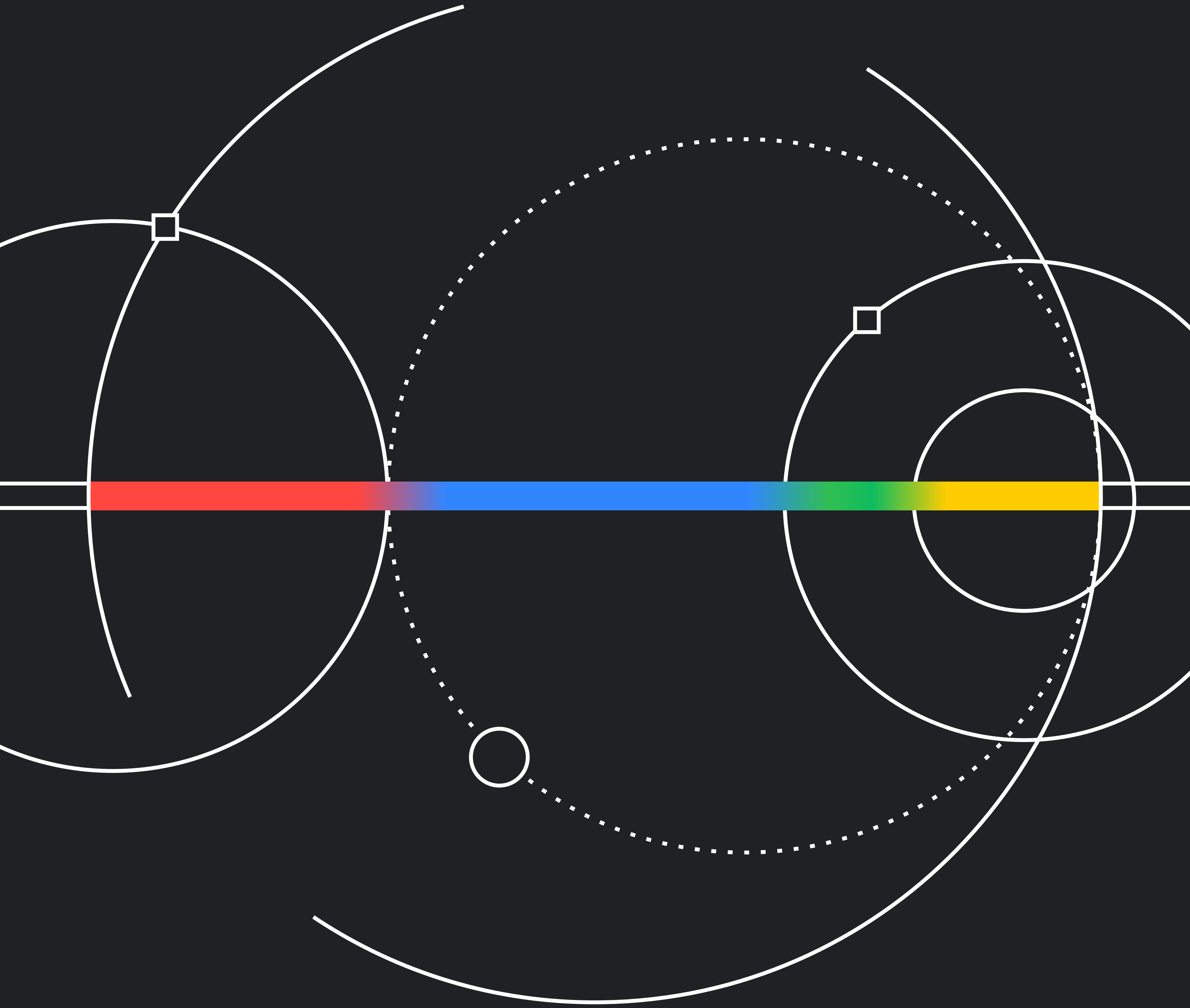
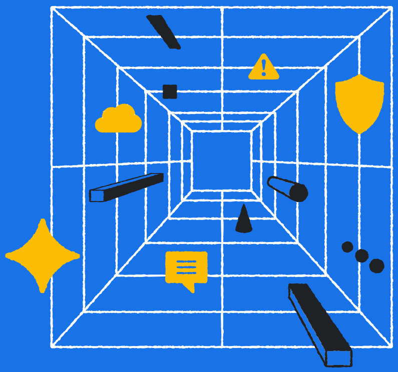


Table of Contents

Introduction

When sharing our insights on cybersecurity in the year ahead, we never make “crystal ball” predictions. Instead, we focus on the real-world trends and data we are observing right now, to provide clear, realistic expectations on what will likely be the biggest trends and challenges.



Our Cybersecurity Forecast 2026 report focuses on three key themes: adversary and defender use of artificial intelligence, cybercrime as the most disruptive global threat, and continued operations by nation state actors to achieve their strategic goals.

Insights were gathered from security leaders across Google Cloud, including Sandra Joyce, VP of Google Threat Intelligence, Charles Carmakal, Chief Technology Officer of Mandiant Consulting, and Jon Ramsey, VP & GM of Google Cloud Security.

The report also features expertise from dozens of researchers, analysts, responders and experts across numerous Google Cloud security teams, including Google Threat Intelligence Group, Mandiant Consulting, Google Security Operations, and Google Cloud’s Office of the CISO. This unique, integrated frontline visibility—from Mandiant’s incident response to Google’s global threat intelligence—allows us to provide a comprehensive forecast of the threats and trends that matter most.

Technology advances, threats evolve, the cybersecurity landscape changes, and defenders must adapt to it all if they want to keep up. The Google Cloud Cybersecurity Forecast 2026 report aims to help the cybersecurity industry frame its fight against cyber adversaries in 2026.

Artificial Intelligence

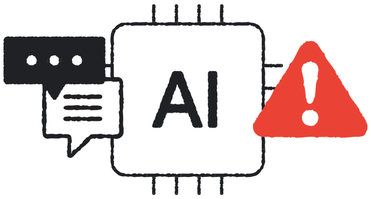
Adversaries Fully Embrace AI

> **“Organizations need to** be prepared for threats and adversaries leveraging artificial intelligence.”
>
> Jon Ramsey, VP & GM of Google Cloud Security

In 2026 and beyond, threat actor use of AI is expected to transition decisively from the exception to the norm, noticeably transforming the cyber threat landscape. We anticipate that actors will fully leverage AI to enhance the speed, scope, and effectiveness of operations, building upon the robust evidence and novel use cases observed in 2025. This includes social engineering, information operations, and malware development.

Additionally, we anticipate threat actors will increasingly adopt agentic systems to streamline and scale attacks by automating steps across the attack lifecycle. We may also begin to see other AI threats increasingly being discussed in security research, such as prompt injection, and direct targeting of the models themselves.

Prompt Injection Manipulates AI



While AI promises unprecedented growth, it also introduces new, sophisticated risks. One of the most critical is prompt injection, a cyberattack that essentially manipulates AI, making it bypass its security protocols and follow an attacker’s hidden command. This isn’t just a future threat; it’s a present danger, and we anticipate a significant rise in these attacks throughout 2026.

The increasing accessibility of powerful AI models and the growing number of businesses integrating them into daily operations create perfect conditions for prompt injection attacks. Threat actors are rapidly refining their techniques, and the low-cost, high-reward nature of these attacks makes them an attractive option. We anticipate a rise in targeted attacks on enterprise AI systems in 2026, as attackers move from proof-of-concept exploits to large-scale data exfiltration and sabotage campaigns.

Google continues to take steps to defend against prompt injection. This involves a multi-layered defense-in-depth approach that includes model hardening, along with system-level guardrails. These guardrails feature machine learning content classifiers to filter malicious instructions from untrusted data, security thought reinforcement to keep the model focused on user intent, and strict output sanitization and user confirmation for high-risk actions.

AI-Enabled Social Engineering

In 2026, we anticipate sophisticated threat actors like ShinyHunters (UNC6240) will accelerate the use of highly manipulative AI-enabled social engineering, making it a significant threat. The key to their success in 2025 was avoiding technical exploits and instead focusing on human weaknesses, particularly through voice phishing (vishing). Vishing is poised to incorporate AI-driven voice cloning to create hyperrealistic impersonations, notably of executives or IT staff.

This approach will be exacerbated by the increasing use of AI in other aspects of social engineering, which threat actors have been leveraging extensively since 2024. This includes reconnaissance, background research, and the crafting of realistic phishing messages. AI allows for scalable, customized attacks that bypass traditional security tools, as the focus is on human weaknesses rather than the technology stack.

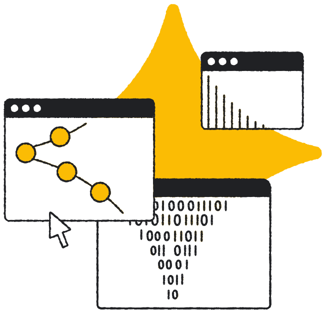
Given the huge success of these social engineering campaigns and the difficulty in apprehending the actors at a deterrent scale, the risk-reward ratio will continue to favor the attackers. Consequently, we expect an increased volume of these attacks in 2026. Defenders must urgently implement processes with multiple checks and balances to defend against these tactics.

AI Agent Paradigm Shift

In 2026, we anticipate the rapid adoption of AI agents for executing workflows and decisions will introduce new challenges, since traditional security deployments were not designed to be operated by AI agents. Organizations will be required to develop and implement comprehensive methodologies, frameworks, and tools to effectively map their new AI ecosystems—and to assess any security vulnerabilities that are introduced.

A central pillar of this new security paradigm will be the evolution of identity and access management (IAM). The concept of identity will expand to treat AI agents as distinct digital actors, each with its own managed identity. This shift will necessitate a move beyond conventional human authentication and service account management towards more dynamic and granular control. We anticipate the rise of "agentic identity management," which will feature adaptive, AI-driven systems for continuous risk evaluation and context-aware access adjustments. The goal is to minimize the potential for privilege creep and unauthorized or unsafe actions. These identity solutions will follow the existing principles of least privilege and will also involve the implementation of just-in-time access, granting temporary and task-specific permissions, as well as a robust chain of delegation.

Supercharged Security Analysts



By 2026, enterprise-wide AI adoption will have fundamentally reshaped the security analyst's day-to-day focus areas. We expect to move past the model of analysts drowning in alerts, and into one where they direct AI agents. This comes as the "Agentic SOC," where frontline intelligence effectively becomes the brain for these new AI partners. For an incident responder, this means an alert will come packaged with a full, AI-generated case summary, a decoded view of that obfuscated PowerShell command, and its mapping to the MITRE ATT&CK framework. The analyst's job shifts from manual data correlation to strategic validation, letting them approve a SOAR containment action in minutes, not hours.

This same principle extends directly to threat hunting and intelligence production. A hunter will be able to form a hypothesis and ask their AI agent in plain English, "Hunt for TTPs (tactics, techniques and procedures) related to UNC5221 across our environment and report anomalies." The AI will perform the heavy lifting of gathering and correlating petabytes of data. An intelligence analyst will be able to provide a malware sample and preliminary notes, tasking the AI to draft a full threat report, complete with actor attribution and mitigations. The AI handles the rote work, letting the analyst focus on the high-level analysis and final judgment. It's about scaling human intuition, not replacing it.

Shadow Agent Risk

By 2026, we expect the proliferation of sophisticated AI Agents will escalate the "Shadow AI" problem into a critical "Shadow Agent" challenge. In organizations, employees will independently deploy these powerful, autonomous agents for work tasks, regardless of corporate approval. This will create invisible, uncontrolled pipelines for sensitive data, potentially leading to data leaks, compliance violations, and IP theft.

Banning agents is not a viable option, as it only drives usage off the corporate network, eliminating visibility. The forward-looking strategy will be to establish a new discipline of AI security and governance. This demands a secure-by-design approach, integrating protection from the start. Companies must deploy AI controls to safely route and monitor all agent traffic. Successful organizations will create working environments that allow for AI innovation while maintaining auditable security.

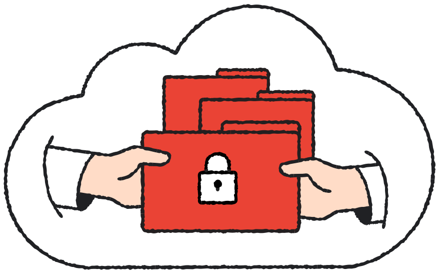
Cybercrime

Ransomware and Data Theft Extortion

> **“We expect to see more** ransomware and extortion. This problem is going to continue and increase in 2026.”
>
> Sandra Joyce, VP of Google Threat Intelligence

In 2026, the combination of ransomware, data theft, and multifaceted extortion will remain the most financially disruptive category of cybercrime globally. This is due not only to the sustained quantity of incidents, but also to the cascading economic fallout that consistently impacts suppliers, customers, and communities beyond the initial victim. For context, incidents in 2025 targeting critical points in retail and food wholesale supply chains resulted in hundreds of millions of dollars in total damages, significantly disrupting consumer supply.

The sheer volume of this activity continues to escalate. The 2,302 victims listed on data leak sites (DLS) in Q1 2025 represented the highest single quarter count observed since we began tracking these sites in 2020, confirming the maturity of the cyber extortion ecosystem. This scale is primarily achieved through major groups' specialized tactics, which include targeting third-party providers and exploiting zero-day vulnerabilities. Targeting managed file transfer (MFT) software allows cybercriminals to execute high-volume data exfiltration across hundreds of targets simultaneously.



Moving into 2026, cybercriminals will continue to utilize initial access strategies such as voice phishing (vishing) and other targeted social engineering techniques to bypass multi-factor authentication (MFA). They may increasingly leverage zero-day vulnerabilities as part of more widespread extortion campaigns, and will also increasingly find more creative ways to coerce victims into paying extortion demands.

The On-Chain Cybercrime Economy

As the financial sector adopts cryptocurrencies and tokenized assets, and moves towards a global on-chain economy, we anticipate threat actors will exploit blockchain characteristics like immutability and decentralization for considerable financial gain. The widespread adoption of crypto and stablecoins rapidly expands the attack surface for both traditional institutions and startups, creating new vulnerabilities in crypto-native solutions and enterprise IT systems alike.

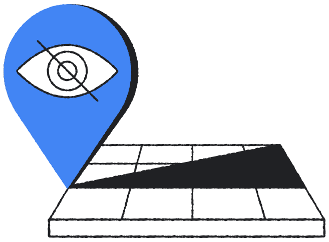
We foresee continued high-value targeting of decentralized finance (DeFi) platforms and cryptocurrency exchanges, including large-scale attacks, and supply chain attacks combined with digital asset theft. Attacks will also continue against regions demonstrating a favorable regulatory stance and growing industry presence, such as the U.S., Southeast Asia, and the Middle East.

Over the next few years, we may begin to see malicious operations migrate core components of their lifecycle onto public blockchains. This could go beyond simple payload delivery using techniques such as EtherHiding. We could see use of the full Web3 stack for a variety of activity, including dynamic command-and-control (C2), decentralized data exfiltration, and asset monetization via tokenized marketplaces. By moving operations on-chain, adversaries gain unprecedented resilience against traditional takedown efforts.

This shift will demand an evolution in defense. The analysts and investigators of 2026 will need to become proficient blockchain investigators, requiring new competencies in tracing transaction histories, decoding malicious smart contract logic, and performing wallet analysis. Organizations that neglect to upskill their teams in these Web3 fundamentals will be blind to an entire class of agile, persistent threat activity.

However, the immutability that grants resilience to the adversary is also a permanent operational security risk for them. Every on-chain action—funding a wallet or deploying a contract—leaves a permanent, publicly auditable record. This will revolutionize attribution, allowing campaigns separated by years to be definitively linked using reused wallet addresses or similar contract bytecode, shifting the fight to the strategic disruption of entire on-chain criminal enterprises.

Enterprise Virtualization Under Threat



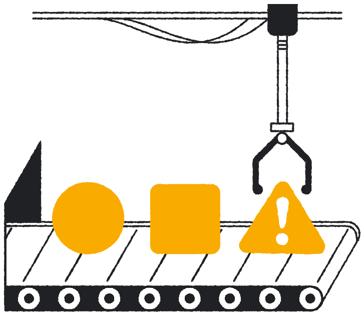
As security controls mature within guest operating systems, we anticipate a significant pivot in threat actor focus towards the underlying virtualization infrastructure—notably for financial gain. This foundational layer, long considered a pillar of strength, is now emerging as a critical blind spot due to a confluence of systemic vulnerabilities: the inherent lack of endpoint detection and response (EDR) visibility, the persistence of outdated software versions, and the prevalence of insecure default configurations. While security teams have concentrated on user endpoints and in-guest defenses, the core virtualization fabric—the host of all enterprise applications—remains largely unmonitored. When combined with deep-seated integrations into legacy core identity services, the hypervisor transforms from an infrastructure component into a high-leverage entry point, where a single compromise can grant adversaries control over the entire digital estate.

This strategic shift by adversaries is not speculative; it represents a tactical evolution toward targets offering maximum impact with minimal resistance. Attacks targeting the hypervisor are designed for systemic disruption. Bypassing in-guest EDR, they execute mass encryption of foundational virtual machine disks, crippling control planes and inducing enterprise-wide operational paralysis. The velocity of this attack vector is a defining factor; adversaries can render hundreds of systems inoperable in a matter of hours, a stark contrast to traditional endpoint ransomware campaigns that often propagate across a network over days or even weeks.

The consequences of a breach at this level are therefore an order of magnitude greater; not just in scale, but also in the radically compressed timeframe for detection and response. Looking forward, safeguarding this often-neglected layer will demand a strategic shift in security strategy, compelling organizations to move beyond guest-centric models and develop new capabilities to counter this escalating threat directly at the infrastructure level.

Google continues to take steps to defend against this threat. Google Cloud VMware Engine (GCVE) enhances security by leveraging managed services and restricting direct access to underlying components like ESXi. This approach offloads security responsibilities to Google, who continuously monitors for vulnerabilities, limiting the attack surface compared to self-managed solutions.

ICS and OT Targeting



In 2026, we anticipate the primary disruptive threat to industrial control systems (ICS) and operational technology (OT) will remain cybercrime. We expect to see ransomware operations specifically designed to impact critical enterprise software (such as ERP systems), severely disrupting the supply chain of data essential for OT operations. This vector is effective because compromising the business layer cripples the industrial environment, forcing quick payments. Meanwhile, poor hygiene like insecure remote access will continue to allow common Windows malware to breach OT networks. Targeted nation-state attacks, though less frequent, will remain highly sophisticated and tied directly to specific geopolitical conflicts.

Defenders will have to prioritize network segmentation to rigorously isolate the OT from the IT network, preventing ransomware from pivoting from the enterprise side. All remote access must be secured with multi-factor authentication (MFA) and least privilege principles to block entry via compromised credentials. To ensure recovery, implement immutable, offline backups of both industrial configurations and critical enterprise data (like ERP logs), and network monitoring to critical IT/OT paths.

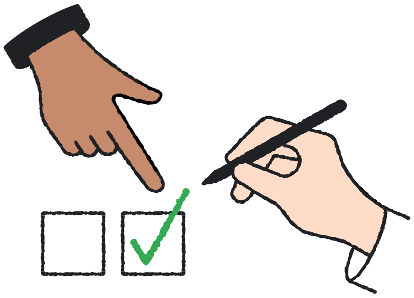
Nation States

Russia

> **“Nation-state adversaries** will continue to penetrate organizations and remain within victim environments for large periods of time.”
>
> Charles Carmakal, Chief Technology Officer of Mandiant Consulting

In 2026 and beyond, Russia’s cyber operations are expected to undergo a strategic shift, moving past a singular focus on short-term tactical support for the conflict in Ukraine to prioritize long-term global strategic goals. While sustained cyber espionage targeting the Ukrainian government and defense sectors will remain a priority—likely seeking critical intelligence for kinetic operations or political developments such as potential peace talks—the apparatus’ focus will widen.

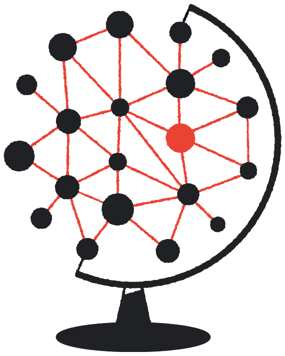
The steady pace of cyber espionage in Europe and North America in 2025, alongside renewed use of novel and creative tactics, techniques and procedures, suggest a transition towards long-term development of advanced cyber capabilities, intelligence collection to support Russia’s global political and economic interests, and obtaining strategic footholds within international critical infrastructure environments. Despite a decline in disruptive and destructive cyberattacks since 2022, organizations must continue to remain vigilant against this threat in 2026.



Pro-Russian information operations (IO) are likely to intensify against the U.S. and other Western nations, while continuing to target Russia’s near abroad. Elections will remain a prime target, as seen in activities related to polls in Poland, Germany, Canada, and Moldova in 2025. Furthermore, IO campaigns will actively seek to manipulate narratives related to news developments, such as promoting claims of alleged Western interference after Romania nullified its 2024 Presidential election.

Pro-Russia hacktivist groups will continue to pose a substantial and unpredictable threat, notably to operational technology (OT) environments, as demonstrated by an April 2025 compromise of a Norwegian dam.

China



In 2026, the volume of China-nexus cyber operations is expected to continue surpassing that of other nations. This sustained, high-pace threat activity will continue to support China's longstanding strategic interests of maintaining internal stability and strengthening its political and economic influence globally. China's cyber threat apparatus is expected to not only maintain its current high volume, but it will also prioritize the ability to conduct stealthy operations and field novel capabilities in the coming year.

We anticipate China-nexus cyber espionage tactics, techniques, and procedures will continue to focus on maximizing operational scale and success, with some threat actors also working to minimize opportunities for detection. China-nexus threat actors will continue to aggressively target edge devices (which typically lack endpoint detection and response solutions), and exploit zero-day vulnerabilities. They will also target third-party providers, since compromising one trusted partner may enable access to many downstream organizations, and abuse of legitimate partner connections makes the resulting malicious access challenging to identify.

One area of particular interest for these operations will be the semiconductor sector, where competition, U.S. export restrictions, and increased demand related to AI adoption may result in espionage, underscoring the importance of a layered approach to network defense.

Concurrently, pro-PRC information operations are expected to continue efforts to shape global perceptions to align with China's strategic interests, specifically by portraying China positively, while negatively framing the U.S., Taiwan, Japan, South Korea, Vietnam, the Philippines, and critics of China.

Iran

Objectives of regime stability and maintaining regional influence amid ongoing geopolitical conflicts are expected to drive Iranian cyber activity in 2026. The escalating regional tensions—exemplified by the Gaza conflict and the exchange of strikes between Iran, Israel, and the U.S. in 2025—will continue to fuel increased cyber espionage, disruptive attacks, and information operations (IO) targeting Israel and its allies.

We anticipate Iranian cyber capabilities will continue to be resilient, multifaceted, and semi-deniable, deliberately blurring the lines between espionage, disruption, hacktivism, and financially motivated activity. This integrated approach allows the same actors and accesses to be leveraged for different missions, complicating defense and attribution for adversaries. Additionally, we anticipate the risk of wiper deployment to remain elevated, building on the aggressive tactics observed since October 2023.

Iran-aligned IO will remain critical for galvanizing support around Middle East conflicts, sowing discord in target countries, and influencing elections. This activity will rely heavily on inauthentic, news-focused websites to seed political content aligned with Tehran's interests. The use of AI-generated content and the amplification of narratives through clusters of inauthentic social media personas, with an increased focus on platforms like Telegram, will accelerate. The ability to rapidly pivot pre-positioned influence infrastructure—as demonstrated by shifts in messaging after the April 2025 Pahalgam Terror Attack—confirms their agility in leveraging emerging global stressors.

Core objectives are expected to remain fixed: continued monitoring of regime critics, intelligence gathering against entities and individuals linked to Iranian or regional politics, and the targeting of technologies that could support the military.

North Korea

North Korea's cyber threat apparatus is expected to sustain their primary objectives of revenue generation and traditional cyber espionage against perceived adversaries, primarily the United States and South Korea, in 2026.

North Korean cyber threat actors will escalate their highly successful and lucrative operations against cryptocurrency organizations and users. The tactics observed in 2025, which included the largest recorded cryptocurrency heist valued at approximately $1.5 billion, provide a clear indication of their focus on high-yield financially motivated attacks. We anticipate that North Korean actors will intensify their technical innovation. This includes tactics such as convincing targets to execute malicious code, and conducting extensive internal reconnaissance of cloud environments to locate and steal high-value assets.

Campaigns will additionally rely on advanced social engineering, such as luring targets with fake "hiring assessment" webpages. Similarly, deepfake videos will become more prevalent to build trust and deceive high-value personnel.

North Korean IT worker activity is projected to continue its expansion globally (notably in Europe), adapting to and attempting to circumvent increased law enforcement pressure and growing awareness in the U.S. This global diversification is a direct reaction to the successful disruption of "laptop farms" in the U.S. that were enabling remote access and obfuscating the workers' true locations.

Furthermore, the risk associated with North Korean IT worker activity will continue to extend beyond simple salary earnings. One objective will be direct financial gain through the abuse of employer network access, specifically targeting and stealing cryptocurrency from crypto-focused organizations. Additionally, workers will continue to leverage their employment access for strategic espionage, as shown by the theft of sensitive data from a defense contractor developing AI technology.

Conclusion

2026 will usher in a new era of AI and security, both for adversaries and defenders. While threat actors will leverage AI to escalate the speed, scope, and effectiveness of attacks, defenders will also harness AI agents to supercharge security operations and enhance analyst capabilities. However, this transformation introduces new challenges such as "Shadow Agent" risks, and the need for evolving identity and access management.

Financially motivated operations, particularly ransomware and data theft extortion, will remain a dominant and disruptive force. Geopolitically, nation-state activity from Russia, China, Iran and North Korea will continue to pose significant and evolving threats, driven by distinct strategic interests and employing diverse cyber tactics.

To navigate this complex and rapidly evolving environment, organizations must prioritize proactive, multi-layered defense strategies, invest in AI governance, and continuously adapt their security postures to safeguard against emerging threats and ensure operational resilience.

The Cybersecurity Forecast 2026 report provides organizations with the essential insights and knowledge needed to navigate the complex threat landscape in the year ahead. By clearly outlining evolving trends and potential threats, this report equips leaders to move beyond reactive defense and build a more resilient, forward-looking security posture.

Contributors

The Cybersecurity Forecast 2026 report features insights from our security leaders:

Sandra Joyce,
VP of Google Threat Intelligence

Charles Carmakal,
Chief Technology Officer of Mandiant Consulting

Jon Ramsey,
VP & GM of Google Cloud Security

Dozens of researchers, analysts, responders and experts from various Google Cloud security teams contributed to this report:

Josh Atkins
Bhavana Bhinder
Doug Bienstock
Sarah Bock
Pierre-Marc Bureau
Michelle Cantos
Stuart Carrera
Anton Chuvakin
Tom Curry
Odun Fadahunsi
David Grout
Adrian Hernandez
Jose Hernandez
Scott Henderson
Joshua Kim
Martin Lawther
Steve Ledzian
Yihao Lim
Keith Lunden
Mark Magee
David Mainor
Stuart McKenzie
Thiébaut Meyer
Jordan Nuce
Josh Palatucci
Christiane Peters
Fred Plan
Alice Revelli
Gabby Roncone
Cameron Sabel
James Sadowski
Nick Schroeder
Chris Sistrunk
Genevieve Stark
Kelli Vanderlee
Alden Wahlstrom
Jess Xia

Google Cloud

For more information, visit cloud.google.com