Introduction

When thinking about the year ahead, some use the term “predictions.” However, our thoughts on the cybersecurity landscape in the coming year have always been based on the trends we are already seeing, and so we feel “forecast” more accurately captures our intentions.

The Google Cloud Cybersecurity Forecast 2024 report is filled with forward-looking thoughts from several of Google Cloud’s security leaders, and dozens of experts across numerous security teams, including Mandiant Intelligence, Mandiant Consulting, Chronicle Security Operations, Google Cloud’s Office of the CISO, and VirusTotal. These individuals are regularly on the frontlines of the latest and largest attacks, and know what organizations and security teams need to be thinking about in the coming year.

Technology advances, threats evolve, attackers change their tactics, techniques and procedures (TTPs), and defenders must adapt if they want to keep up. The Google Cloud Cybersecurity Forecast 2024 report aims to help the cybersecurity industry frame its fight against cyber adversaries in 2024.
Improved, professionalized, and scaled phishing

Generative AI and large language models (LLMs) will be utilized in phishing, SMS, and other social engineering operations to make the content and material (including voice and video) appear more legitimate. Misspellings, grammar errors, and lack of cultural context will be harder to spot in phishing emails and messages. LLMs will be able to translate and clean up translations too, making it even harder for users to spot phishing based on the verbiage itself. LLMs will allow an attacker to feed in legitimate content, and generate a modified version that looks, flows, and reads like the original, but suits the goals of the attacker.

With gen AI, attackers will also be able to execute these campaigns at scale. If an attacker has access to names, organizations, job titles, departments, or even health data, they can now target a large set of people with very personal, tailored, convincing emails. A malicious LLM may not even be necessary to create these emails since there is nothing inherently malicious about, for example, using gen AI to draft an invoice reminder.

Scalable information operations

A clever gen AI prompt will be all attackers need to create fake news, fake phone calls that will actively interact with recipients, and deepfake photos and videos based on gen AI-created fake content. These operations could increasingly enter into the mainstream news cycle. With scalability of these types of information operations comes the risk of reducing public trust in news and (online) information, to the point where everyone will become more skeptical—or simply stop trusting—what they see and read. This could make it increasingly difficult for businesses and governments to engage with their audiences in the near future.

We judge that such gen AI technologies have the potential to significantly augment information operations—and other operations such as intrusions—in the future, enabling threat actors with limited resources and capabilities, similar to the advantages provided by exploit frameworks such as Metasploit or Cobalt Strike. Adversaries are already experimenting with gen AI, and we expect to see more use of these tools over time.
Interpreting data, understanding threats, and shoring up defenses

Cyber defenders will use gen AI and related technologies to strengthen detection, response, and attribution of adversaries at scale, as well as speed up analysis and other time-consuming tasks such as reverse engineering. A big use case of AI is to drive how organizations will synthesize large amounts of data, and contextualize it in threat intelligence to then yield actionable detections or other analysis. We will see this come to fruition in 2024, with AI and gen AI providing the ability to augment human capability in analyzing and inferring actions to take from these large data sets. We will see new ways of overlaying customer specific data in a highly confidential way, giving organizations the ability to take significant action at speed and scale. This will be one of the bigger transformations for organizations leveraging AI for security purposes in the coming years, ultimately helping them to reduce toil, address threat overload, and close the widening talent gap.

Gen AI and LLMs as a service... for attacks

LLMs and other gen AI tools will increasingly be developed and offered as a service to assist attackers with target compromises. They will be offered in underground forums as a paid service, and used for various purposes such as phishing campaigns and spreading disinformation. We’ve already seen attackers have success with other underground as a service offerings, including ransomware used in cyber crime operations.
The Big Four

China

Activity from China will continue to be driven by long term priorities such as internal stability and territorial integrity, including issues related to Taiwan, Chinese regional hegemony and influence, and economic influence over key markets. Chinese cyber espionage actors will continue to preserve stealth, reduce opportunities for detection, and stymie attribution. We expect to see continued use of tactics such as zero-day exploitation, targeting of systems on the network edge, supply chain compromise, and botnets and proxy networks designed to disguise traffic both within a compromised network and between the threat actors and a victim.

Additionally, China is expected to continue development of a military and civilian force capable of launching disruptive and destructive operations, and campaigns in support of national political and military objectives. The potential for disruptive and destructive operations to be carried out by Chinese threat actors during times of active conflict poses a threat to organizations globally, and could impact essential daily life activities, critical infrastructure, and safety.

Russia

We expect Ukraine to remain a primary focus of Russian cyber threat activity in 2024 and beyond, with intelligence gathering, disruptive and destructive attacks, and information operations occurring at elevated rates. We will also continue to observe Russian cyber espionage operations—likely strategic intelligence gathering missions—outside of Ukraine that are consistent with longstanding priorities, including targeting of government, defense, civil society and non-profits, and energy.

Sanctions on Russia will continue to hurt technological and military innovation in the country. Russia will likely resort to increased intellectual property theft to compensate for lack of domestic expertise. This behavior will be modeled after Chinese intellectual property theft that has occurred over the last few years.
North Korea

North Korea-based cyber threat activity has had an increased emphasis on financially motivated operations, notably targeting the cryptocurrency industry as well as other blockchain-related platforms. In 2024, we expect North Korea to place even heavier emphasis on stealing cryptocurrency—and NFTs—to fund their weapons and nuclear program in addition to enabling their cyber operations and infrastructure acquisition.

Notably, we have observed the country run self-sustaining operations to diminish the financial strain on North Korea’s central governing bodies. This approach to funding aligns with the regime’s ideology of juche, or self-reliance leading to collective prosperity. In recent years we observed a relative increase in cyber crime campaigns being used to fund espionage operations, and expect that trend to continue. We also expect North Korea will take advantage of opportunities to perform more supply chain compromises.

Iran

We expect that Iran’s geopolitical ambitions, economic development needs, competition with regional rivals Saudi Arabia and Israel, threats to regime stability and survivability, and surveillance of the Iranian diaspora and opposition groups will be key drivers of state-sponsored cyber threat activity in the coming year.

We believe cyber espionage actors associated with Iran, as well as Palestinian and Lebanese threat actors, present an increased threat to Israel following Hamas’ multi-pronged kinetic assault on civilian and military targets in central and southern Israel on Oct. 7, 2023. We anticipate that Iranian threat actors are likely to conduct intelligence gathering, information operations, and potentially hybrid hack-and-leak or other disruptive and destructive attacks.
Global forecasts

Continued use of zero-day vulnerabilities (and edge devices)

We have observed a general increase in zero-day vulnerability use since 2012, and 2023 is on track to beat the previous record set in 2021. We expect to see more zero-day use in 2024 by both nation-state attackers as well as cyber criminal groups. One of the reasons for this is that attackers want to maintain persistent access to the environment for as long as possible, and by exploiting zero-day vulnerabilities (as well as edge devices), they’re able to maintain access to an environment for much longer than if they were to, for example, send a phishing email and then deploy malware. Security teams and solutions have become much better at identifying malicious phishing emails and malware, and so attackers will turn to other avenues to evade detection. Edge devices and virtualization software are particularly attractive to threat actors because they are challenging to monitor. For cyber criminals, they know using a zero-day vulnerability will increase the number of victims and, based on recent mass extortion events, the number of organizations that may pay high ransomware or extortion demands.

Cyber activity targeting U.S. elections

As we move into a United States presidential election year, we will see nation states and other threat actors engage in a variety of cyber activity, including espionage and influence operations targeting the electoral systems, impersonation of candidates on social media, and information operations designed to target the voters themselves. We also don’t expect operations to decrease following the elections. We will likely see an uptick in spear phishing and other attacks against the U.S. government as nation states—particularly China, Russia and Iran—seek to gain a decision advantage (potentially during an administration change). These campaigns may feel more prevalent in 2024, as gen AI tools are leveraged to increase scale and operational tempo.
Rise of disruptive hacktivism

In 2022 and 2023, we observed a resurgence in the volume of hacktivist activity, particularly associated with threat actors expressing support for Russia or Ukraine during the ongoing Russian invasion. Likewise, recent conflict between Hamas and Israel has been accompanied by a flurry of hacktivist activity. Observed activity includes distributed denial-of-service (DDoS) attacks, data leaks, and defacements. Notably, in both contexts Mandiant Intelligence is tracking ostensible hacktivist groups that exhibit some greater than average capabilities, and significant alignment with narratives and objectives of the state they claim to support. While we cannot confirm ties to nation state operators at this time, we note that Russian and Iranian groups have used false hacktivist fronts in the past. Mandiant Intelligence judges that the perceived success of such operations as providing sufficient plausible deniability increases the likelihood that states will use such cyber attacks against civilian and military targets, and we surmise that this could extend to the use of these tactics to achieve kinetic damage.

Wipers become a standard capability in all nation state cyber arsenals

Before the 2022 Russian invasion of Ukraine, Russian APT groups gained access to Ukrainian targets and launched a destructive attack that coincided with kinetic operations. Other nation states will copy this technique by adding wiper malware to their cyber arsenals. With tensions in the Taiwan Strait and other global security threats, 2024 will likely see pre-placement of destructive wiper malware at strategically important targets.

Targeting of space-based infrastructure

The situation in Ukraine has demonstrated dependencies on space-based technologies (Starlink, for example, and other satellites and communications networks) during conflict. In 2024, we expect to see evidence of sophisticated state-sponsored cyber actors’ full spectrum Computer Network Exploitation capabilities to compromise space-based and associated ground support infrastructure and communications channels to interdict, disrupt, deny, degrade, destroy, or deceive an adversary—as well as to conduct espionage.
Attacks targeting hybrid and multicloud environments mature and become more impactful

In 2023, Mandiant worked with VMware to remediate a zero-day vulnerability that allowed the attacker to execute code on guest virtual machines (VMs). Even though the impact of this vulnerability was limited to a single hypervisor, it proved that threat actors were targeting cloud environments looking for ways to establish persistence and move laterally. In 2024, we will see these techniques evolve to cross boundaries between cloud environments. Threat actors will look to exploit misconfigurations and identity issues to move laterally across different cloud environments.

Serverless services in the cloud more heavily used by threat actors

In 2023, we saw an increase in cryptominers being deployed on serverless infrastructure. In 2024, we predict that cyber criminals and nation-state cyber operators will more heavily leverage serverless technologies within the cloud. Attackers will move towards serverless for the same reasons developers are adopting serverless; they offer greater scalability, flexibility, and can be deployed using automated tools.

Extortion operations continue

Extortion operations remain likely the most impactful form of cyber crime to enterprises and societies worldwide. Despite a stagnation in growth during 2022, advertisements for stolen data and extortion revenue estimates indicate that this threat is growing in 2023, and we anticipate this growth will continue in 2024 without a significant, market-wide disruption.
Espionage and “sleeper botnets”

Cyber espionage operations will continue to find more ways to scale their attacks while also creating additional OPSEC for their operations. Espionage groups will create “sleeper botnets” out of vulnerable Internet of Things, small office, home office (SOHO), and end of life devices and routers using a mixture of old and new exploits. These “sleeper botnets” will be used as needed, and discarded once caught or no longer useful, complicating efforts to track and attribute activity. These “sleeper botnets” will differ from traditional botnets where the number of devices were used to amplify attacks, such as DDoS attacks.

Revival of ancient techniques

While attackers are incorporating new techniques to evade detection, we expect to see some actors resurrect ancient techniques that aren’t widely covered. For example, in 2013, a researcher wrote a blog post about using undocumented SystemFunctionXXX functions instead of cryptography functions in the documented Windows API. This technique didn’t become popular until Q4 2022, when several security researchers began discussing it and releasing code snippets in their own blogs and on GitHub. At that point, more malware samples implementing this technique started popping up on VirusTotal. We also observed recent use of an anti-virtual machine (anti-VM) technique detailed in a 2012 malware analysis book. The technique was not covered in detection rules because the hypervisor is not often used in many countries.

Continued migration to modern programming languages by malware authors

Malware authors will continue to develop more software in programming languages such as Go, Rust, and Swift. This is because the languages provide a great development experience, low level capabilities, large standard library, and easy integration with third-party packages. These languages and ecosystems enable rapid development of complex malware, making it cheaper to write new malware to evade detection. This means a churn in toolsets used by actors, and a corresponding need for new detection signatures. Unfortunately, these modern languages often bring a large runtime (Go) and/or use the latest compiler techniques (Rust) that make reverse engineering tasks more difficult. In other words, the benefits of packing and obfuscation without using a protector.
Developers targeted in supply chain attacks via software package managers

In recent years, supply chain attacks against NPM (the Node.js package manager) such as IconBurst have demonstrated how threat actors target software developers. In one particularly concerning scenario, a developer is compromised by installing a malicious package, which gives a threat actor access to the developer’s source code, and allows the actor to add a backdoor. This is a low-cost, high-impact attack. As a result, the prevalence of these sorts of attacks is likely to continue to grow, particularly as threat actors shift to other, less monitored package managers such as PyPI (Python) and crates.io (Rust). We need to remain vigilant in monitoring these sources of software libraries.

Growing prevalence of mobile cyber crime

In 2024, we anticipate cyber criminals or scammers to continue employing novel social engineering tactics such as simulating domestic help services, messages from fake social media accounts, banks or government officials, and spoofed pop-up alerts to trick victims into installing malicious applications on their mobile devices.

Cyber insurance premiums remain steady

The insurance market is known for fluctuations. A hard market indicates that premiums are rising and coverage is restricting, while a soft market indicates that premiums are decreasing and coverage is broadening. After a strong correction in the cyber insurance market over the past few years that was characterized by rising premiums and restricted coverage, the market is starting to soften. With more entrants in the market and insurers with ambitious cyber growth goals, the competition is expected to provide much needed relief to the rising premiums the industry has been seeing. While we expect to continue to see a general trend towards restrictions in systemic risk coverage, it’s possible that insurers may broaden coverage in other ways to compete in this new landscape.
Consolidation around SecOps

In 2024, we expect to see more consolidation in SecOps as customers increasingly demand integrated risk and threat intelligence in their security operations solutions. Customers are going to demand an integrated ecosystem that covers their entire network estate—cloud, multicloud, on-premises, and hybrid environments—and will increasingly expect vendors to offer opinionated workflows, guidance, and content to jumpstart their security program out of the box.
Cyber activity around elections

In 2024, Taiwan, South Korea, India, and Indonesia are some of the countries that will be holding elections. We have previously observed cyber espionage, cyber crime, hacktivism, and information operations actors express interest in these pivotal events. We anticipate observing election lures being leveraged for scams, as well as intelligence gathering purposes. China’s newly drawn out map could also become a cause of contention during India’s and Indonesia’s elections.

“Pig butchering” scams to be an ongoing problem

Pig butchering scams, which have elements of both cyber crime and human trafficking, will continue to be a problem in 2024 for JAPAC countries’ law enforcement. Pig butchering scams are a type of online fraud in which scammers pose as potential romantic partners over extended periods of time in order to gain the trust of their victims. Once they have gained their victim’s trust, the scammers will begin to convince them to invest in various fraudulent financial schemes. An August 2023 UN report detailed that many of these scammers are themselves victims, and have been trafficked and forced to work in scamming operations. In July 2023, 2,700 people were rescued from forced cyber crime labor in the Philippines. The UN report claims, “the situation remains fluid: hundreds of thousands of people from across the region and beyond have been forcibly engaged in online criminality.”
Shifting tactics, techniques and procedures

Endpoint detection and response solutions are becoming more widespread in the JAPAC region, and overall organizations are becoming more security mature. As a result, well-resourced threat actors will increasingly leverage tactics intended to minimize opportunities for detection. We are already seeing this globally. Defenders in the region should prepare for exploitation of zero-days in security, networking, and virtualization software; targeting of routers and other edge devices; and use of other methods to relay and disguise attacker traffic both outside and inside victim networks.
Disinformation campaigns in Africa

In the digital age, disinformation has become a powerful tool for geopolitical influence. Russia and China are increasingly targeting African countries with cyber campaigns designed to spread misinformation, sow discord, and undermine democratic institutions, and we do not see them slowing down in 2024. We expect to see Chinese and Russian groups targeting the rare earth minerals industry since they are essential for many high-tech products such as smartphones, computers, and electric vehicles. By gaining control of these resources, Russia and China can strengthen their economic and strategic positions in Africa.

Another way that Russia and China will use disinformation to influence Africa is by supporting authoritarian regimes. These regimes often crack down on dissent and restrict access to information. This makes it easier for Russia and China to spread their propaganda and undermine democratic values. Disinformation and focus on Africa is a long-run game, and we see 2024 as a year of peak activity on this front.
Olympics 2024 broadens the attack surface in Paris (and beyond)

During the 2024 Summer Olympics in Paris, we expect to see cyber criminals targeting ticketing systems and merchandise, particularly through a surge in phishing campaigns requesting financial information or credentials. Public authorities and banks need to remain vigilant. We may also see geopolitical activity that involves using the Olympics to try to destabilize and put pressure on France, and through it Europe and the style of political regime associated with it. The Olympics will also likely be a target for misinformation and disinformation, whether directly linked to the event (ticket sales, for example) or indirectly linked to it (accommodation rentals, public transport).
Conclusion

While new technologies will aid security teams, they can also expand the attack surface. In 2024, the rapidly evolving world of gen AI will provide attackers with new ways to conduct convincing phishing campaigns and information operations at scale. However, defenders will use the same technologies to strengthen detection, response, and attribution of adversaries—and more broadly reduce toil, address threat overload, and close the widening skills gap.

Next year we expect to see continued activity by The Big Four—China, Russia, North Korea, and Iran—as they conduct espionage, cyber crime, information operations, and other campaigns to achieve their individual goals. Since organizations are becoming better at security, many of these attacks will involve techniques to evade detection, including use of zero-day vulnerabilities and the targeting of edge devices.

Everyone should be prepared for global activity around the myriad major events being held throughout 2024, including the U.S., European Parliament and other elections, as well as the Summer Olympics in Paris. Additionally, as major global conflicts continue into next year, be prepared for an uptick in disruptive hacktivism.

The cybersecurity landscape is constantly evolving, sometimes in new and unexpected ways. Defenders, often with limited resources, have the monumental task of keeping up. The Google Cloud Cybersecurity Forecast 2024 report is our way of helping security professionals prepare for the certainties and uncertainties of the year ahead. We hope our knowledge from the frontlines helps you feel ready.
Contributors

Our Cybersecurity Forecast 2024 report features insights from Google Cloud’s security leaders, including:

Charles Carmakal, CTO of Mandiant Consulting
Sandra Joyce, VP of Mandiant Intelligence
Sunil Potti, GM and VP of Cloud Security
Phil Venables, Chief Information Security Officer

Many others across Google Cloud also contributed to the report:

Willi Ballenthin  Mike Hom  Mike Raggi
Dan Black  Renze Jongman  Alice Revelli
Sarah Bock  Dan Kennedy  Nick Richard
Anton Chuvakin  Cris Kittner  Matt Shelton
Jamie Collier  Karen Kukoda  Monica Shokrai
Vivek Chudgar  Steve Ledzian  Daniel Sislo
Charles deBeck  Yihao Lim  Genevieve Stark
Vicente Diaz  Keith Lunden  Kelli Vanderlee
Eric Doerr  Jens Monrad  Alden Wahlstrom
Renato Fontana  Joseph Pisano  Dominik Weber
David Grout  Fred Plan  Richard Weiss
Scott Henderson  Ofir Rozmann  Jess Xia