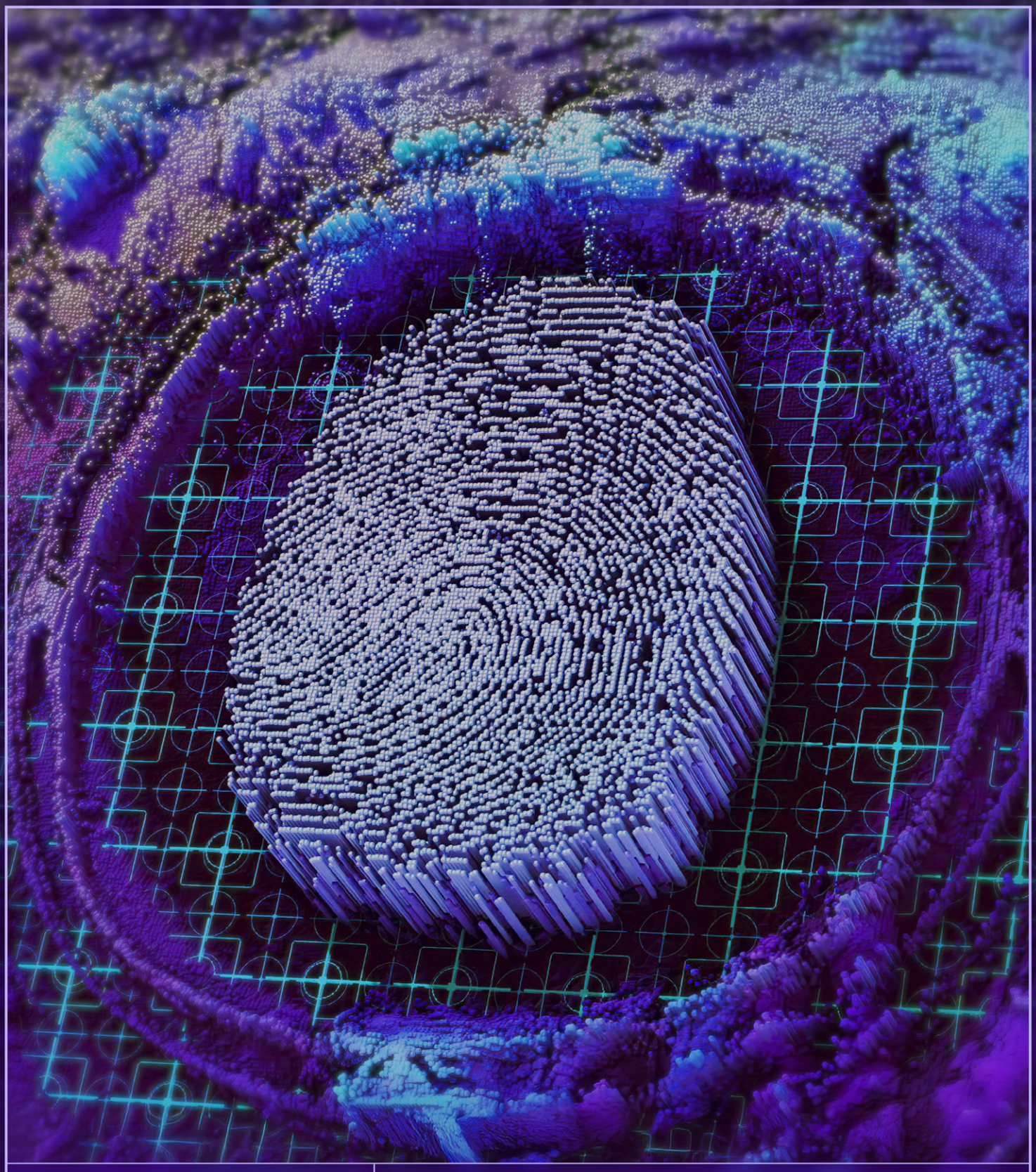
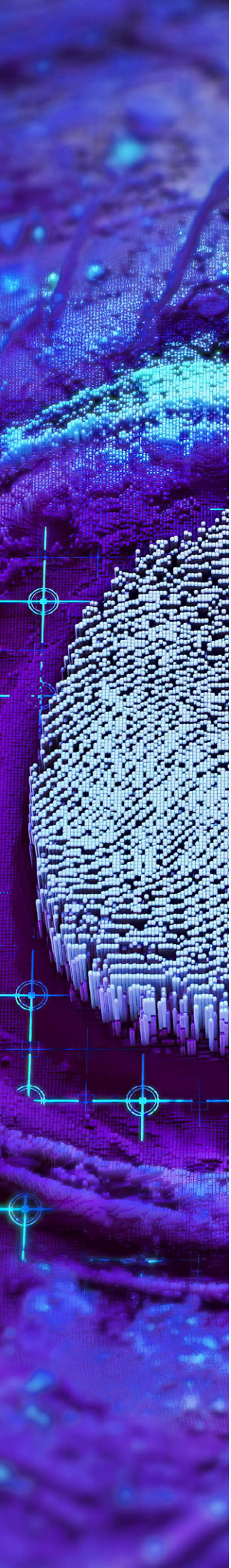


Ontinue

2 H 2 0 2 5

Threat Intelligence Report



Executive Summary

The second half of 2025 made one trend unmistakably clear: **attackers aren't breaking in anymore, they're logging in**. Identity compromise, cloud credential abuse, and the exploitation of trust in SaaS and automation pipelines have become the defining characteristics of today's threat landscape. Ransomware, phishing, and supply chain attacks haven't gone away, they've simply hitched a ride on this identity driven model.

Think of identity as the *new perimeter*, but also the *new skeleton key*. Once attackers get valid credentials (human or not), they can slip through the front door without tripping the alarms traditionally meant to catch intruders. And in 2H 2025, the market for those keys became industrialized.

This report was developed by Ontinue's Advanced Threat Operations (ATO) team, with contributions from Ontinue threat researchers, incident responders, and security analysts. Drawing on investigations conducted across customer environments and telemetry from the Ontinue ION SecOps platform, the report analyzes the techniques, campaigns, and threat actor behaviors observed during the second half of 2025.

The Big Shift: Identity as the Primary Attack Surface

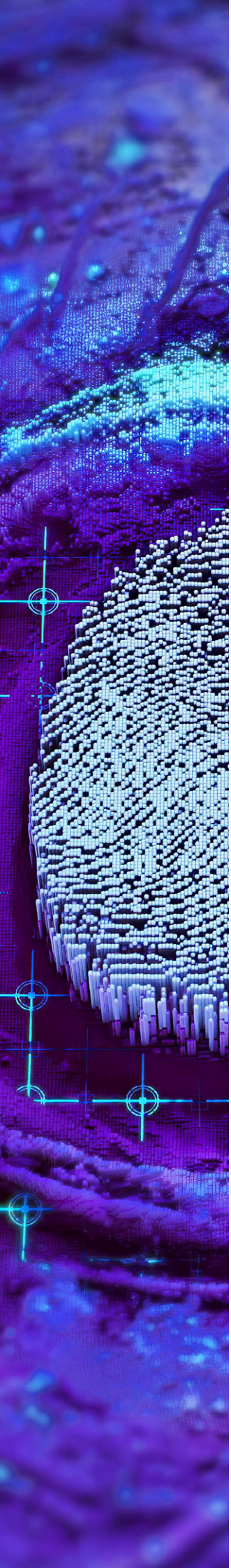
Across our telemetry and investigations, identity based attacks dominated true positives. Password sprays, adversary in the middle (AiTM) phishing kits, OAuth token abuse, and leaked service principal credentials became the most common entry points for compromise. Attackers increasingly rely on the fact that authentication systems are easier to manipulate than hardened endpoints.

The surge in infostealer activity drove this shift. Tools like Lumma, RisePro, RedLine, and Vidar fueled an underground ecosystem where stolen credentials, session tokens, and cloud console access were openly sold - sometimes for thousands of dollars. It's credential theft at industrial scale.

Ransomware Evolves, Again

Ransomware remained one of the most disruptive threats of 2025, with more than 7,000 incidents and 129 active groups tracked. However, the real evolution wasn't in encryption, it was in the extortion strategy. Modern ransomware attacks blend encryption, data theft, operational disruption, and direct pressure campaigns into layered extortion models.

Even though ransom payments dipped slightly year-over-year ($892M in 2024 to $820M in 2025), the operational and economic costs continue to climb, amplified by the identity driven compromises that enable attackers to move faster and with more precision.



Infostealers: The Hidden Engine Behind Modern Cybercrime

Infostealers became the quiet workhorses of the criminal economy with small tools creating big problems. Delivered through fake installers, malvertising, phishing, archives, or malicious npm packages, they siphon everything from browser cookies to cloud tokens. This data then fuels ransomware, fraud, and access as a service operations.

The market for these tools is fully tiered with low end utilities available for $10–$50/month, mid-tier for a few hundred, and premium offerings for up to $1,000/month. It's a SaaS model optimized for crime.

Node.js based infostealers also rose in prominence. By bundling malicious runtimes with AI themed or productivity apps, campaigns like TamperedChef made credential theft feel like just another software install.

Supply Chain & SaaS: Trust as an Attack Vector

Supply chain attacks accelerated, leveraging automation pipelines and non human identities. Two incidents stood out:

- The **Salesloft OAuth abuse campaign**, which affected 700+ organizations by exploiting trusted app integrations.
- The **Shai Hulud npm worm**, which republished up to 100 compromised packages per victim, weaponizing developer workflows at scale.

In both cases, attackers didn't need zero days, they simply abused trust relationships embedded deep in automation, CI/CD processes, and OAuth permissions.

Cloud Service Principals: Automation's New Weak Point

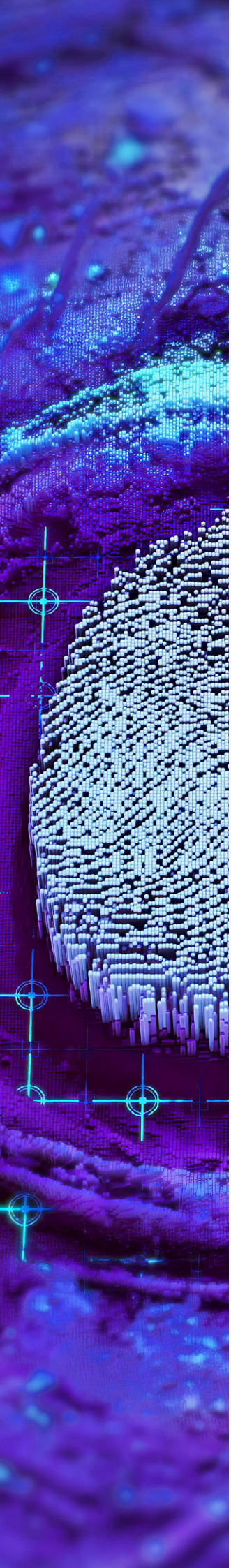
Service principal credentials - often stored in config files or build artifacts - became a high value target. Tools like TruffleHog made it trivial for attackers to scan, validate, and exploit these secrets. Once compromised, service principals enable seamless, MFA-less access to cloud APIs, resource enumeration, and long term persistence.

Since their activity often resembles normal automation, detection is slow, giving attackers more room to operate.

AI Enters the Attacker Toolkit

Generative AI's impact on malware development became visible in 2H 2025. Ontinue's threat researchers observed telltale signs of LLM assisted coding, including overly verbose comments, code duplication, and polished UIs with insecure logic.

These early indicators suggest that as AI enhanced tooling becomes more accessible, the barrier to entry for sophisticated attacks will continue to drop.



Geopolitical Pressure Rises, and Civilian Targets Feel It

State aligned actors remained active and aggressive. Highlights included:

- The Lazarus Group's **$1.5B cryptocurrency theft**
- Wiper attacks targeting **Polish civilian infrastructure** by Ghost Blizzard
- Record setting DDoS activity peaking at **31.4 Tbps** via botnets with more than 500,000 IPs

Civilian entities - SaaS companies, energy providers, even individual users - were increasingly treated as acceptable collateral in geopolitical campaigns.

What This Means for 2026

Across every trend, one message repeats: **identity is now the central battlefield of cybersecurity**. Human identities, machine identities, service principals, OAuth granted apps, and CI/CD tokens all represent potential front doors to your environment.

Organizations that thrive in 2026 will be those that treat identity security not as a project, but as a continuous discipline, supported by phish-resistant authentication, strong governance, and monitoring tailored to both human and non-human access.

CHAPTER 1

The Identity Centric Threat Landscape

Identity has become the place where modern attacks truly take shape. In 2H 2025, attackers no longer treated identity as a side channel. They treated it as the control plane that governs everything else. When someone can get a valid token or a working pair of credentials, they don't need to "break in" at all. They just blend in. This section outlines how identity became the dominant attack surface and why defenders must think of identity the same way earlier generations thought about the perimeter.

1.1 From Perimeter to Identity Perimeter

The era of network-centric and device-centric boundaries is over. Across our telemetry, the most reliable indicators of active compromise weren't exploit attempts or malware infections - they were authentication events. Real threats announced themselves through password sprays, AiTM phishing flows using kits such as Tycoon2FA and Salty2FA, OAuth token abuse, and the exposure of service principal credentials.

Attackers increasingly behave like sophisticated users rather than system intruders. They focus on login pathways because identity systems still allow them to impersonate users cleanly. Once authenticated, even improperly, they inherit the trust the organization has given to that identity. The reality is reveals an uncomfortable truth: identity is the new perimeter because it's the only boundary attackers must cross to behave like insiders.

1.2 Industrialized Credential Theft and the Access Economy

Credential theft is no longer a byproduct of malware, it's a business model. In 2H 2025 we saw credential markets expand in both scale and sophistication. Listings tied to LummaC2 alone surged by 72%, with high privilege cloud console credentials selling for $1,000–$15,000+.

This growth was driven by infostealers that harvested everything from browser cookies to MFA bypass artifacts. These logs flowed into an underground "access economy" that fuels ransomware, supply chain attacks, and fraud. Infostealers don't need to be complex; they just need to capture the right data. And once that data is exfiltrated, attackers can rapidly turn it into authenticated cloud sessions, often without touching your network at all.

Think of this market like a digital pawn shop. It's operationally cheap to run, stocked by thousands of infections per day, and visited by adversaries looking for the path of least resistance.

1.3 Non-Human Identities: Cloud's Most Overlooked Risk

While human credentials drive headlines, non-human identities drive compromises. Service principals, like automation accounts, CI/CD identities, and background application entities, were one of the most consistently mishandled identity types across environments in 2H 2025.

Our investigations revealed repeated patterns: secrets like ClientID and ClientSecret stored in plaintext configuration files, deployment artifacts, or public repos; leaked tokens found through automated secret scanners such as TruffleHog; and attackers authenticating directly to Microsoft cloud APIs using OAuth2 client credential flows.

The danger lies in how easily attackers achieve persistence. Service principals often have broad permissions, operate without MFA, and generate traffic that looks like normal automation. As a result, defenders detect these intrusions late—long after enumeration or privilege escalation has occurred.

1.4 Where Authentication Breaks: AiTM and Session Hijacking

Adversary in the middle techniques surged, thanks to increasingly polished phishing kits like Tycoon2FA and Salty2FA. These kits captured credentials and session cookies in real time, turning MFA into an obstacle easily bypassed rather than a hardened layer of defense. Attackers used infrastructure protections such as Cloudflare Turnstile to mask malicious pages, making them difficult for SOC analysts to validate during investigations.

AiTM is effective because it reframes identity theft: instead of cracking a lock, attackers simply forward authentication traffic and collect the keys. They borrow the user's session rather than the password alone. This isn't a failure of MFA;it's a failure of relying on authentication events that attackers can proxy.

1.5 Identity Abuse as the Trigger for Supply Chain and SaaS Compromise

Identity compromise doesn't stop at user accounts, it cascades outward. Two notable cases in 2H 2025 encapsulate this risk:

- **Salesloft OAuth Abuse**: attackers exploited trusted OAuth connections, impacting more than 700 organizations.
- **Shai Hulud npm Worm**: attackers used stolen developer tokens to republish up to 100 compromised packages per victim, weaponizing automation pipelines into distribution channels.

These incidents show that once attackers compromise identity - especially automation identities - they can exploit trust relationships embedded deeply in SaaS apps, DevOps workflows, and CI/CD systems. Authentication becomes the attacker's passport into a supply chain that was never designed to verify intent.

1.6 Identity in Detection: The New Epicenter of True Positives

This trend was born out in Ontinue's Cyber Defense Center. In 2H 2025, identity driven events dominated our high fidelity detections. Password spray attempts, anomalous login flows, suspicious OAuth grants, and service principal misuse consistently surfaced as top-tier alerts. Although the signatures varied, the underlying theme was consistent: if authentication looked strange, adversaries were usually involved.

This trend also extended into hybrid attack chains. For example, investigations frequently linked DGA-based communication to identity compromise, with attackers mixing reconnaissance and credential misuse in the same operation. Even when malware appeared, it was often a precursor or companion to a broader identity takeover effort.

1.7 Why Identity Has Become the Boundary That Matters

Across every cybersecurity trend, we've seen that identity has become the primary battlefield of cybersecurity. The data is clear that things likepassword-based compromises, token theft, session hijacking, OAuth abuse, and API-key exposure all represent attacker-preferred vectors. Identity is attractive because it grants authenticity, not just access.

In short, the attacker no longer needs to "break" anything when they can simply log in with the same privileges as your users, apps, and automation systems.

CHAPTER 2

Deep Dive: Major Threat Trends of 2H 2025

The threat landscape of 2H 2025 wasn't defined by one "big" adversary or one new technique. Instead, it was shaped by the acceleration of several parallel forces such as, industrialized credential theft, the rise of identity driven cloud compromise, increasingly professionalized infostealer operations, rapidly evolving ransomware tradecraft, the mainstreaming of 2FA bypassing phishing kits, and the first meaningful signs of AI assisted malware development.

If chapter 1 describes the what, this chapter is the how. It examines the threat trends that consistently drove real world incidents, not theoretical risks, but patterns repeatedly seen by Ontinue's ATO and CDC teams.

2.1 Ransomware: High Volume, Strategic Evolution

Ransomware remained one of the most disruptive trends of the period, with more than **7,000 claimed incidents** across 2025 and at least **129 active groups** contributing to the ecosystem. Even as cryptocurrency payments declined slightly (from $892M in 2024 to $820M in 2025), the operational impact on victims continued to grow.

What changed isn't the encryption, it's the business model. Ransomware operators now run layered extortion:

- **Encryption** (classic ransomware)
  - Business impact manifests as downtime and recovery costs
  - Leverage: Companies must pay for decryption / keys and occasionally 'support'
- **Data theft**
  - Attackers exfiltrate sensitive data first, then threaten to publish it on a leak site or sell it
  - Leverage: shifts from IT recovery to legal/regulatory exposure, customer trust, IP loss, and contractual penalties
- **Operational disruption**
  - Beyond encrypting endpoints, attackers disrupt the ability to operate by wiping systems, deleting backups, sabotaging virtualization, attacking OT/ICS-adjacent services, or breaking identity/administration planes
  - Leverage: Even if you restore, you'll be down longer and it will cost more
- **Direct pressure on employees, customers, or partners**
  - Attackers contact employees, executives, customers, partners, or the media to increase urgency and reputational damage
  - This can include threats of notifying regulators/customers, or drip-feeding "proof" data releases to show they're serious
  - Leverage: turns a private incident into a public/time-sensitive crisis

Think of modern ransomware as a multi-layer extortion machine. Even when victims avoid paying, they are still dealing with downtime, regulatory exposure, third party disruption, and long recovery cycles. This model is effective because it:

- **Reduces defender options**: backups only solve on layer (encryption), not theft and/or exposure
- **Shifts pain to leadership**: legal, PR, and revenue risk often outweigh IT concerns
- **Creates a countdown**: harassment and staged leaks manufacture urgency
- **Makes non-payment costly**: attackers try to ensure you still pay even if recovery is possible

2.1.1 Why ransomware thrives in an identity-first world

Most ransomware operations begin not with exploits but with purchased or stolen credentials. Infostealer logs, cloud console access, and non-human identity misuse give attackers clean initial access, allowing them to move laterally far faster than in traditional malware-led intrusions.

2.2 Infostealers: The Engine of the Credential Economy

If ransomware is the visible crisis, infostealers are the unseen machinery that fuels it. In 2H 2025, infostealers remained the most consistently impactful malware category across customer incidents. These tools generally fall into major "families", each with their own tactics, ecosystem, and preferred targets. While their names change frequently, a small number of dominant strains account for most real-world incidents because they are inexpensive, fast to operate, and easy to deploy. They include:

Lumma Stealer (Lumma C2)

- A well-known stealer as a service that grew rapidly and has been the biggest threat of 2025
- Known for evasion techniques and frequent updates
- Targets browser data, crypto wallets, and authentication tokens
- Widely spread via fake installers and malicious ads

RedLine

- A stealer that surged through malvertising campaigns
- Often delivered via fake software installers
- Steals browser data, crypto wallets, and system info
- Frequently found in logs sold on marketplaces

RisePro

- One of the most widespread stealers on the market
- Sold on underground forums as a subscription
- Steals browser passwords, cookies, autofill, crypto wallets
- Often delivered through fake installers, cracks, and malvertising
- Frequently bundled with loaders like PrivateLoader

Vidar

- A modular stealer with strong persistence features.
- Often delivered via malvertising and fake software downloads
- Can download additional payloads
- Steals credentials, browser data, crypto wallets, and Telegram sessions

Raccoon

- A long running stealer with multiple versions
- Known for extremely fast data exfiltration
- Distributed via phishing, fake software, and exploit kits
- Steals credentials, cookies, crypto wallets, and system info

Stealc

- A highly customizable stealer sold on a subscription model
- Steals credentials, cookies, crypto wallets, and messaging tokens
- Delivered via loaders, phishing, and cracked software
- Known for its modular plugin system

Other notable stealers include:

Rhadamanthys

- A complex, multi-modular malware often used in "ClickFix" campaigns that trick users into running malicious code.

MetaStealer (META)

- A cross-platform infostealer that has gained popularity for targeting macOS systems in addition to Windows.

Formbook/xLoader

- A persistent threat known for keylogging and stealing form data.

2.2.1 Why infostealers were so effective this period

Infostealers succeeded because they took advantage of how people naturally work:

- Users install "free" utilities, PDF tools, productivity apps, and AI-themed software.
- Attackers hide payloads inside these installers.
- Within minutes, browser cookies, passwords, and MFA session tokens are exfiltrated.
- These logs feed a thriving underground market — with cloud access priced at $1,000–$15,000+.

The market matured as well. As previously mentioned, LummaC2 listings alone grew 72%, highlighting the industrialization of access resale.

2.2.2 Node.js Infostealers: Trust as a Delivery Channel

A noteworthy subset of cases involved trojanized installers bundling a full Node.js runtime, a clever trick that let attackers run malicious JavaScript inside a familiar, legitimate framework. These loaders quietly executed scripts to steal credentials, establish persistence, and communicate with attacker infrastructure. One example of this observed in the wild is TamperedChef.

It's the equivalent of hiding in plain sight: attackers using the target's own tools and execution layers against them.

Compared to campaigns that beacon to obvious "bulletproof hosting," Node.js infostealer activity often uses newly registered domains plus mainstream cloud/CDN hosting to make C2 and payload traffic look like normal SaaS and web delivery. Cloud delivery also enables fast rotation of domains and endpoints without major disruption, complicating reputation-based controls and static blocklists making detection challenging.

2.3 Cloud Service Principal Abuse: Identity Without a Human in the Loop

One of the most strategically important themes of 2H 2025 was the rise in attacks targeting non-human identities, especially cloud service principals. These identities often hold elevated privileges and do not use MFA, making them an ideal target.

2.3.1 Common attack chain observed

1. **Secret leakage**
   Plaintext ClientID/ClientSecret pairs appeared in appsettings.json files across developer machines, web servers, and CI/CD pipelines.
2. **Automated discovery**
   Tools like TruffleHog scanned repos, validated leaked secrets, and attempted authentication.
3. **OAuth token acquisition**
   Attackers used legitimate Azure AD OAuth2 flows to obtain application-only tokens.
4. **Graph API enumeration**
   Directory reconnaissance followed, with broad 404 patterns suggesting probing.
5. **Persistence / Data Access**
   Attackers accessed files, enumerated directory roles, and used the app's permissions to maintain stealthy footholds.

What makes this alarming is how "normal" the traffic looks. These attacks use legitimate APIs, legitimate flows, and legitimate credentials, they simply belong to the wrong person.

2.4 Shai-Hulud: Identity-Driven Supply Chain Worm

Another prime example of non-human identity exploitation manifests in the compromise of CI/CD pipelines and developer supply chains. Shai Hulud is an aggressive, npm focused supply chain worm that weaponizes everything modern development pipelines were designed to streamline automation, trust, and speed. While earlier versions were already capable of covert propagation through malicious npm packages, the latest Shai Hulud version significantly upgrades its timing, stealth, and ability to hijack the CI/CD ecosystem itself.

Rather than slow rolling its way through the supply chain like XZ or SolarWinds, Shai Hulud 2.0 is fast, opportunistic, and ruthlessly efficient, spreading through maintainers' own identities and tooling.

2.4.1 Primary Targets

Shai Hulud focuses on the soft underbelly of software supply chains. It targets:

- **Small npm maintainers** responsible for widely used dependencies but lack hardened workflows
- **Automated CI/CD environments**, especially GitHub Actions runners, where pre install scripts run with high implicit trust
- **Development ecosystems** that rapidly auto ingest "latest" package releases without cooling off periods
- **Organizations with exposed tokens** in build logs, config files, or ephemeral runner environments

These targets give the worm an ideal blend of ubiquity, trust, and automation to spread with minimal resistance.

2.4.2 Notable Capabilities

The latest Shai Hulud variant doesn't rely on exotic exploits, it abuses what developers already consider "normal."

Pre Install Pipeline Hijack

The latest version of Shai Hulud introduces a preinstall hook script (setup_bun.js / set_bun.js) that triggers before tests, scanners, or code reviews run in GitHub Actions or similar CI systems. This is its most important evolution.

What it does:

- Fetches **Bun**, a fast JavaScript runtime
- Executes a **"fat," obfuscated second stage (bun_environment.js)**
- Runs entirely on ephemeral CI/CD runners, leaving almost no post job artifacts

This timing shift allows the malware to outrun nearly all early pipeline detection.

Credential Harvesting via Deceptively Legitimate Tools

Shai Hulud scans for:

- npm tokens
- GitHub PATs
- Cloud credentials
- Local secret stores and metadata endpoints

It uses **TruffleHog**, a tool normally associated with good security hygiene. CI logs showing TruffleHog don't raise suspicion, giving the malware near perfect camouflage while it quietly scoops up your secrets.

Covert Exfiltration Using Victims' Own Infrastructure

Shai Hulud exfiltrates stolen secrets to new public GitHub repositories, sometimes created under *another victim's* identity. It hides data inside double or triple encoded blobs, labeled with cheeky descriptors such as "Sha1 Hulud: The Second Coming."

This multi-victim piggybacking makes it significantly harder to follow the breadcrumbs.

High-Speed, Worm Style Propagation
Using the tokens it steals, the worm immediately:

- Bumps your package versions
- Republishes them under your name
- Pushes out up to ~100 compromised packages per victim

Your account becomes an unwilling amplifier, broadcasting malware to your entire dependency graph.

Malicious Workflow Injection & Rogue Runners (New Behavior)

A major change in the updated version is its ability to:

- Plant malicious GitHub workflows for later execution
- Register fraudulent self hosted runners, often named SHA1HULUD
- Use these runners as footholds to harvest fresh secrets long after the initial infection

This is a significant escalation. The malware is no longer just a package hijacker, it's a CI/CD persistence mechanism.

Destructive Failover Logic
If Shai Hulud finds little to harvest and no clear path to propagate, version 2 may attempt to delete the user's home directory.

This nihilistic "scorched earth" fallback is new and signals the author's willingness to cause irreversible damage.

2.4.3 Why It Works Now

Modern development pipelines reward speed, trust, and automation, which arepreciselythe factors Shai Hulud abuses.

- Developer package installations generate heavy "expected noise," masking malicious activity
- CI/CD pipelines implicitly trust dependency updates and execute arbitrary scripts
- Automated runners allow ephemeral, unlogged execution environments
- Small maintainers often lack monitoring and are highly attractive targets
- In a landscape obsessed with "move fast," Shai Hulud thrives by moving even faster.

2.5 Phishing Kits: 2FA Bypass Becomes a Commodity

Adversary in the middle AiTM phishing surged in 2H 2025. Kits like **Tycoon2FA**, **Salty2FA**, and **Evilginx** no longer required technical sophistication. They came as polished, subscription-based toolkits.

2.5.1 What has changed

Attackers began systematically blocking datacenter IPs. Analysts loading a phishing URL from corporate networks would see a clean page, while employees at home saw a perfect impersonation of Microsoft 365 or Google Workspace.
Cloudflare features like the below were consistently abused to hide malicious content from investigators and prevent takedown.

- **Turnstile CAPTCHA** – Used to filter automated analysis tools and sandboxes before serving phishing content
- **IP Geofencing** – Datacentre IP ranges are systematically blocked, returning benign content or errors to analysts while serving malicious pages to residential IPs
- **DDoS protection** – Legitimate Cloudflare features repurposed to protect malicious infrastructure from takedown and analysis

This means SOC teams have to adapt. Residential VPNs, mobile hotspots, and out-of-band inspection became necessary just to validate a suspicious URL.

2.6 GenAI-Assisted Malware: Early Indicators, Real Impact

Ontinue observed the first meaningful signs of LLM-assisted malware development in 2H 2025. This didn't look like autonomous malware, it looked like human developers leaning on AI for speed, features, and UI polish.

2.6.1 What the team observed in real samples

Analysis of a PHP webshell recovered revealed clear indicators of LLM-assisted development. The shell featured polished UI elements (gradient buttons, hover effects, consistent colour theming) alongside fundamental security weaknesses. This pattern is characteristic of developers prompting for features without understanding the full threat model.

Indicators of LLM generation:

- Verbose inline comments explaining obvious code behavior
- Duplicate or inconsistent functions (e.g., mixed variable naming), likely from iterative prompting
- UI components (CSS gradients, styling) unusually polished
- Security “best practices” implemented incorrectly (e.g., bcrypt but no CSRF)

Security vulnerabilities in the shell:

- XSS vulnerabilities
- Path traversal

LLMs didn’t write the malware, but they wrote large pieces of it. This lowers the bar dramatically. Adversaries with minimal engineering ability now ship tools that look more professional but still contain fundamental security flaws.

These threat trends are interconnected. Infostealers feed cloud compromises. Cloud compromises feed ransomware. SaaS and supply-chain attacks exploit identities harvested earlier. AiTM phishing fuels credential theft, which fuels everything else.

Identity is the common thread running through every major trend in 2H 2025.

CHAPTER 3

Geopolitical Forces Reshaping the Cyber Threat Landscape

The second half of 2025 largely continued the geopolitical and cybersecurity patterns established earlier in the year. Governments and critical industries remain exposed as loosely regulated commercial data practices undermine operational security. State-sponsored actors increasingly treat civilian sectors as legitimate targets in ongoing geopolitical competition. At the same time, China continues to demonstrate how long term planning and coordinated national strategy can accelerate progress in strategic technology domains.

3.1 Data Brokers and the Erosion of Operational Security

A recent **investigation by Le Monde** highlighted how **advertising derived location and behavioral data can be used to identify sensitive government personnel**, including home addresses and daily routines, with minimal technical effort. The study also surfaced risks to staff in critical national infrastructure, such as nuclear facilities.

These findings reinforce long standing concerns: **data brokers create exploitable visibility into individuals whose roles should remain opaque**. Organizations operating in sectors tied to national security or high value intellectual property must treat commercial data exposure as a real operational threat. Measures such as restricting personal devices on site, enforcing strict social media policies, and reducing digital footprints materially limit adversaries' ability to map and target key personnel.

3.2 Civilian Industries as Geopolitical Targets

In December 2025, **Poland's CERT reported coordinated wiper malware attacks against civilian heating and energy providers**. Microsoft attributed the activity to the "Ghost Blizzard" threat actor, known for targeting industrial systems.

Although Poland is not engaged in open conflict, the attack targeting critical services for nearly half a million people and represents a clear escalation in hybrid warfare. Civilian entities, regardless of their geopolitical posture, should no longer assume immunity from state linked operations. Organizations must reassess their threat models, understand how their industry aligns with national strategic interests, and prepare for targeted activity even outside traditional conflict zones.

3.3 China's Drive for Semiconductor Self Sufficiency

Ongoing **U.S. export controls have accelerated China's push to build a resilient domestic semiconductor ecosystem**. While China still trails Western and Taiwanese manufacturers in capability and scale, and faces demographic and supply chain challenges, the country continues to leverage long term planning and political commitment to close the gap.

This trajectory illustrates how **strategic constraints can motivate states to pursue asymmetric progress, particularly in sectors foundational to national security.** It also reinforces the broader dynamic: geopolitical competition increasingly concentrates around technology supply chains, innovation velocity, and access to specialized components.

3.4 Conclusion

These developments are unlikely to shift dramatically in the near term, but they set the stage for evolving adversary behavior in 2026. Security and threat intelligence teams should continue to evaluate how their organization fits into nation-level, geopolitical interests, strengthen identity and data centric defenses, and monitor how employee routines or digital footprints may unintentionally create exposure.

CHAPTER 4

In the News

The **Bybit exchange** was hit in February 2025, with **~$1.5 billion in Ethereum stolen by the North Korean Lazarus Group**. Ben Zhou, ByBit confirmed the attack in a post on X (formerly Twitter), explaining that a planned transfer was manipulated, allowing hackers to drain the exchange's ETH cold wallet. More information surfaced after the DPRK-linked malware developer's computer was infected by LummaC2 Infostealer, revealing links to the hack.

In July 2025, Ontinue published **early warning and hunting query to counter the threat dubbed "ToolShell"**, which compromised over 400 Microsoft SharePoint servers worldwide.

One of the **costliest operational cyberattacks of 2025 struck Jaguar Land Rover,** forcing weeks-long production shutdowns and impacting the broader automotive supply chain - estimated economic damage in the billions.

A coordinated **ransomware campaign hit major UK retailers - including Marks & Spencer, Co-op, and Harrods** - exploiting third-party suppliers. Impacts included outages for online sales and critical systems, and heavy financial losses.

Attackers reportedly **exfiltrated millions of SSO and LDAP identity records from Oracle Cloud infrastructure**, potentially affecting hundreds of thousands of tenants with exposed encryption keys and login credentials

In 2025, **ClickFix attacks** became widely adopted by numerous threat actors, including state-sponsored hacking groups and ransomware gangs. What started as a Windows malware campaign, quickly expanded to macOS and Linux, with attacks that installed infostealers, RATs, and other malware.

Clickfix further **evolved with GlitchFix/ErrFix**, users thinking their computers stopped working properly. Different payloads are deployed depending on the victim's host, through the ErrTraffic Traffic Distribution System(TDS).

2025 saw record-breaking distributed denial-of-service (DDoS) attacks targeting organizations worldwide. **Multiple incidents mitigated by Cloudflare demonstrated the increasing firepower of DDoS platforms**, with attacks peaking at 5.6 Tbps, 7.3 Tbps, 11.5 Tbps, and later 22.2 Tbps.

Much of this growth was attributed to the **Aisuru botnet**, which emerged as a significant force behind some of the largest DDoS attacks ever recorded. Microsoft reported that Aisuru leveraged more than 500,000 IP addresses in a 15 Tbps attack targeting Azure, with Cloudflare later reporting that the botnet was responsible for an even larger 29.7 Tbps DDoS attack, peaking again in January 2026 at 31.4 Tbps.

In 2025, **North Korean IT workers infiltrating Western companies became a massive identity threat facing organizations**. The US government says that these workers funnel their earnings to the DPRK regime to fund its weapons program and other initiatives. Rather than exploiting software vulnerabilities, North Korean actors increasingly used fake identities, intermediaries, and legitimate employment to gain access to Western companies, often remaining undetected for long periods.

Software developers have been also continuously targeted by Lazarus in the threat campaign dubbed as Contagious Interview since as early as 2022. The developers are invited to download and execute malware as part of the developer interview. Critical employees are **targeted through fake recruiters and hundreds of domains**, the interviews include increasingly creative ways such as trojanized videoconferencing software installers or errors employing ClickFix techniques, downloading malicious npm packages, and applying backdoors such as BeaverTail or Invisible Ferret.

The 2025 **Salesloft cyberattack was a major SaaS supply-chain incident where attackers abused stolen OAuth tokens from Salesloft's Drift AI Chat Agent integration** to gain trusted access to customers' Salesforce environments. Without exploiting Salesforce, attackers exfiltrated CRM data from 700+ organizations by leveraging legitimate API access. The incident highlighted the growing risk of over-privileged third-party integrations and non-human identities in cloud environments.

Insider threats had a massive impact in 2025, **CrowdStrike disclosed that it detected an insider feeding information to hackers**, including screenshots of internal systems. The insider was reportedly paid $25,000 by a group calling itself the "Scattered Lapsus$ Hunters," a name referring to overlapping threat actors associated with Scattered Spider, Lapsus$, and ShinyHunters.

Our 2024 H1 Report **"The Rising Threat: Chinese State Sponsored Attacks" section warned about broad network exploitation**, which is still in an upward trend. We experience almost monthly sometimes trivial-looking Remote Code Executions (RCEs), Authentication Bypasses, resulting in most organizations may be objectively exposed to active, wide-scale exploitation up to 2-8 times per year depending on their brand of Internet exposed network devices alone.

Meanwhile the use of zero-days has unfortunately become the norm, intelligence sharing about active exploitation – or in cases, infuriatingly even timely fixes – and detailed disclosures are becoming the exception.

The "Joint Cybersecurity Advisory Countering Chinese State-Sponsored Actors Compromise of Networks Worldwide" has been an important attempt to inform, safeguard and monitor the numerous major network security brands - which have been aggressively researched for vulnerabilities in great detail over the past years.

4.1 Disruptions by Law Enforcement and Technology

In 2025 alone, mostly thanks to international Law Enforcement coordination including Operation Endgame, we have **seen over a dozen disruption events, including Qbot, Rhadamantys, Danabot, LummaC2 being staple names among criminals**. Some of these even resulted in Indictments.

Telegram undertook a massive cleanup effort in 2025, removing tens of millions of channels and groups linked to cybercrime, fraud, doxxing/extortion, terrorism, and other illegal activity. Reports estimate that 20–44 million such channels and groups were taken down over the course of the year. This wave of removals was part of a broader shift toward intensified moderation, leveraging user reports, machine learning, and new AI-driven detection systems to reduce illicit activity on the platform.

In addition to removing malicious channels, Telegram also **targeted high profile criminal marketplaces**. Two of the largest illicit crypto market ecosystems, **Huione Guarantee and Xinbi Guarantee, were taken offline in May 2025.** These marketplaces had collectively facilitated more than $35 billion in illicit stablecoin transactions, making their removal one of the most significant blows to cryptocurrency enabled cybercrime that year. These actions underscored a growing emphasis on disrupting the financial infrastructure that supports large scale cybercriminal operations.

4.2 Where Banned Channels and Communities Migrated?

Some underground communities have moved off Telegram altogether to other decentralized or privacy-focused messengers like SimpleX, especially after repeated takedowns. For example, a group called **BFRepo migrated to SimpleX** in May 2025 and rebuilt a presence there.

Groups that have been shut down (e.g., hacktivist or hack forum communities) sometimes resurface on dark web forums or in places like BreachForums, RAMP, X (formerly Twitter) and related sites, particularly after repeated bans.

As Telegram moderation ramps up (with large numbers of illicit cybercrime and doxxing/blackmail channels being blocked), **some underground communities discuss or use alternatives like Tox, Signal, Matrix, and others as fallback communication channels.**

4.3 Indictments and arrests

Major administrators behind prominent cybercrime forums BreachForums and XSS were arrested, significantly disrupting the sale of stolen data and malware kits.

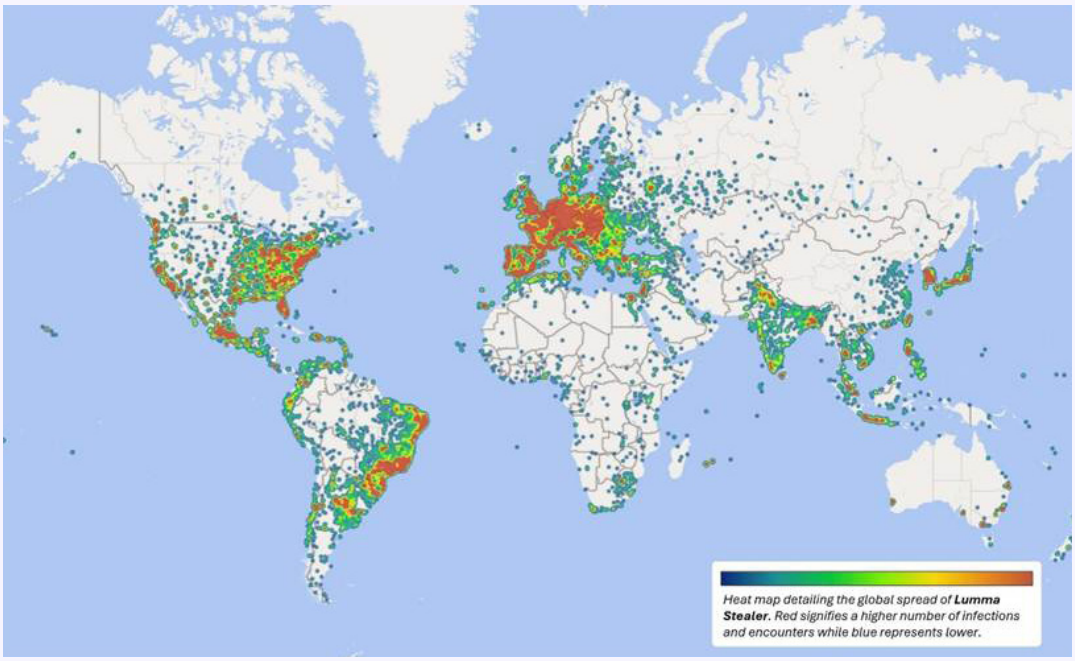
Source: ic3.gov

A major international effort (dubbed Operation Endgame 3.0 around 10–13 November 2025) dismantled cybercriminal infrastructure called the DanaBot malware network, showed a major step preventing further cybercrime totaling $50 million in losses and over 300.000 device infections.

Pro-Russian Hacktivist Group NoName057(16) has been disrupted by Operation Eastwood led by Europol, resulting in two arrests and warrants being issued against 5 core members. NoName recruits a cyber "mercenary army" of ordinary citizens with gamified tasks for a number of years and have been disruptive and targeting primarily Ukraine and various NATO countries in the government, transportation, logistics, finance, energy sectors and major events with pro-Ukraine sentiment.

Source: Europol

Interpol and law enforcement action led to the **disruption of over 2,300 domains acting as C2 servers for the info-stealing malware**, responsible for over 10 million infections. The popular **Infostealer had seen a 72% increase in stolen data listings** in early 2025 prior the disruption, while still operating after the major hit.



Source:https://www.microsoft.com/en-us/security/blog/2025/05/21/lumma-stealer-breaking-down-the-delivery-techniques-and-capabilities-of-a-prolific-infostealer/

The North Korean IT Worker Crackdown, a joint effort to expose a scheme where North Korean operatives posed as freelance IT workers operating from Residential IPs – recruited citizens from Gig economy apps facilitating their "laptop farms" from their homes, which generated revenue for North Korea's nuclear weapons program, as investigated by the FBI's Cyber and Counterintelligence Divisions' "DPRK RevGen: Domestic Enabler Initiative".

In May 2025, the leader of the **Qakbot malware conspiracy was indicted** for involvement in a global ransomware scheme, with $24 million in cryptocurrencies (30 BTC and over $700,000 of USDT) seized.

The U.S. DOJ indicted a Ukrainian national for involvement in the **Nefilim Ransomware campaign** affecting hundreds of companies worldwide.

In December 2025, a former senior manager of a government defense contractor was indicted for fraud, misleading federal agencies about the security of a cloud platform used by the U.S. Army.

Ransomware Negotiation Fraud: A former incident responder was charged with using the ALPHV/BlackCat ransomware variant to launch attacks.

A former US defense contractor employee has pleaded guilty to selling stolen exploit tools to a Russian broker that resulted in estimated $35 million in damages.

In Bavaria (Germany), over 3,000 illegal trading platforms were reported, and more than 100 criminal charges filed, with over €25 million in criminal proceeds seized, as part of efforts targeting fake brokers and cyber investment scams.

A **core member of the notorious Scattered Spider cybercrime gang was sentenced to 10 years in U.S. federal prison and ordered to pay $13 million in restitution** for SIM-swapping schemes and crypto theft. He pleaded guilty to wire fraud and identity theft earlier in the year.

CHAPTER 5

Understanding the Intersection between Identity & Cloud

As discussed throughout this report, identity has become the primary control plane for security. When an identity is compromised, whether through credential theft, token replay, or phishing, attackers can bypass traditional defenses and operate within the environment as legitimate users.

Securing identity at the point where it interacts with cloud resources is therefore critical. Authentication must verify not only who the user claims to be, but also the context in which access is being requested, such as device health, location, and risk level.

By understanding how authentication works across major cloud providers and by implementing strong identity governance, organizations can significantly reduce the risk of unauthorized access and lateral movement.

This chapter provides an overview of the most critical point of control: the authentication layer where identity meets cloud infrastructure. It also offers guidance for how organizations should approach hardening this point of control in their environments.

5.1 Cloud Authentication Mechanisms

Each major cloud provider approaches authentication with a slightly different philosophy, but all share the same objective: validating identity before granting access to sensitive resources.

Microsoft Azure (Microsoft Entra ID)

In Microsoft environments, Conditional Access policies act as the primary gatekeeper for authentication decisions. Conditional Access evaluates multiple factors, including user identity, device state, location, and risk signals, before allowing access to resources.

A common misconfiguration involves using permissive logic when defining authentication requirements. For example, allowing access with MFA OR a compliant device can create unintended bypass paths. A stronger approach is enforcing MFA AND a compliant device, ensuring attackers cannot reuse stolen tokens from unmanaged systems.

Organizations should also regularly audit Conditional Access policies to identify exclusions. Service accounts and legacy authentication exceptions often become blind spots that attackers actively search for.

Amazon Web Services (AWS)
In AWS environments, authentication is primarily managed through IAM Identity Center. The platform emphasizes centralized authentication and encourages the use of temporary credentials rather than long-lived static access keys.

When a user attempts to authenticate, IAM Identity Center evaluates the login request against defined policies, including multi-factor authentication requirements and source attributes. For sensitive operations, organizations can enforce authentication flows that require hardware security keys or additional verification before administrative actions are allowed.

This approach reduces the risk associated with compromised credentials by limiting the lifetime and scope of access tokens.

Google Cloud Platform (GCP)
Google Cloud Platform takes a context-driven approach through Context-Aware Access (CAA). Instead of validating identity alone, CAA evaluates both the user and the environment from which the access request originates.

Even if a user presents valid credentials, access may be denied if the request originates from an unexpected location, an unmanaged device, or a network that does not meet security policy requirements. This combination of identity and contextual validation enables more granular control over access to cloud consoles and APIs.

Continuous session validation further ensures that authentication remains trustworthy throughout the interaction, not just at the moment of login.

5.2 Comparing Authentication Approaches Across Clouds

Although each provider implements identity controls differently, their authentication models share common goals.

Microsoft's Conditional Access acts as a policy-based gatekeeper that evaluates risk signals before granting entry. Google Cloud's Context-Aware Access emphasizes contextual validation, combining identity with environmental attributes such as device posture and location. AWS Identity Center relies on federated authentication and temporary credentials, placing strong emphasis on identity verification and short-lived access tokens.

Despite these differences, all three approaches reinforce the same principle: identity must be continuously validated, not simply verified once.

Feature	Microsoft (Conditional Access)	GCP (Context-Aware Access)	AWS (IAM Identity Center)
Philosophy	**Gatekeeper:** Check then enter	**Holistic:** Identity + Context Combined	**Federated:** Trust the source
Identity Factor	User, Group, Risk Level	Verified User + Attributes	User, MFA Status
Logic	User, Risk, Device State	Identity, IP, Location, Device	Identity Tags, Source IP

5.3 Building a Multi-Cloud Identity Strategy

For organizations operating across multiple cloud platforms, identity architecture plays a decisive role in security. Without a unified approach, environments quickly suffer from identity sprawl - multiple disconnected authentication systems that increase complexity and reduce visibility.

A common best practice is establishing a centralized identity provider as a single source of truth. Many organizations rely on Microsoft Entra ID for this role, using it to federate authentication into AWS and Google Cloud environments.

This model provides a consistent login experience while allowing each cloud platform to enforce its own internal authorization controls. It also simplifies governance by consolidating authentication policies in one location.

Another widely adopted safeguard is the two-account model. Employees maintain one account for everyday productivity tasks and a separate administrative identity for infrastructure management. The administrative account is protected by stronger authentication policies and used only when privileged actions are required.

5.4 The Changing Landscape of Permissions Management

Managing permissions across multi-cloud environments has historically been addressed through Cloud Infrastructure Entitlement Management (CIEM) solutions. Microsoft previously offered this capability through Entra Permissions Management, following its acquisition of CloudKnox.

However, recent changes have shifted these capabilities into the broader Microsoft Defender ecosystem. As the standalone product is phased out, organizations are reassessing how they maintain visibility into cross-cloud permissions.

The core challenge remains understanding who can access what resources, and under what conditions. Many organizations are now implementing identity analytics platforms capable of mapping relationships across identities, workloads, and permissions. These systems create a unified view of identity activity across infrastructure and SaaS environments, helping security teams detect excessive privileges and reduce potential attack paths.

5.5 The Move Toward Passwordless Authentication

One of the most significant identity security developments arriving in 2026 is Microsoft's shift toward passwordless authentication by default. Beginning in April 2026, Microsoft Entra ID will automatically enable passkey profiles for tenants, prioritizing FIDO2-based credentials over traditional passwords.

Passkeys bind authentication to a specific device and cryptographic identity, eliminating many of the weaknesses associated with passwords and traditional MFA methods. Since authentication occurs through cryptographic verification rather than shared secrets, passkeys significantly reduce the effectiveness of phishing attacks and MFA fatigue campaigns.

Organizations should prepare for this transition by educating users about passkey registration and ensuring helpdesk teams are ready to assist with onboarding. While the shift may initially require operational adjustments, it represents a meaningful step toward reducing one of the most persistent attack vectors in modern cybersecurity.

CHAPTER 6

Best Practices for 2026

Building Resilience in an Identity-First Threat Landscape

The trends observed in 2H 2025 reinforce a clear mandate: security programs must evolve beyond perimeter defense and endpoint visibility. Organizations should prioritize the following measures heading into 2026.

Enforce Identity as the Security Control Plane

Identity is now the primary attack surface.

- Require phish-resistant authentication (FIDO2/passkeys) for all privileged users.
- Eliminate Conditional Access exclusions wherever possible.
- Enforce strict "MFA AND compliant device" logic - not OR.
- Separate standard user accounts from administrative identities.

Compromised credentials should not automatically translate into access.

Secure Non-Human Identities

Service principals, OAuth apps, CI/CD tokens, and API integrations are increasingly targeted.

- Remove hardcoded secrets from code repositories.
- Implement secret vaulting and automatic rotation.
- Monitor for anomalous application-based sign-ins.
- Review OAuth permissions and eliminate overprivileged third-party integrations.

Non-human identities must be governed with the same rigor as human accounts.

Reduce Supply Chain Exposure

Development pipelines and SaaS trust relationships are now frontline attack surfaces.

- Implement token expiration and rapid revocation policies.
- Enforce dependency governance and delay auto-ingestion of newly published packages.
- Maintain an up-to-date Software Bill of Materials (SBOM).
- Audit CI/CD workflows for embedded secrets and unusual runner activity.

Speed is the attacker's advantage. Friction can be a defender's strength.

Disrupt the Credential Economy

Infostealers and Access-as-a-Service thrive on poor credential hygiene.

- Deploy endpoint controls that detect browser credential harvesting.
- Revoke active sessions and refresh tokens during password resets.
- Monitor dark web marketplaces for exposed corporate access.
- Implement device-based risk scoring before granting cloud access.

Assume credentials will be harvested. Build controls that invalidate them quickly.

Harden Against Ransomware Through Identity Controls

Modern ransomware frequently begins with purchased access.

- Restrict administrative privileges and enforce just-in-time elevation.
- Monitor for abnormal token usage and privilege escalation patterns.
- Validate backup integrity and isolate recovery infrastructure from domain trust.
- Conduct tabletop exercises that simulate identity compromise, not just malware detonation.

Reducing blast radius is as critical as preventing entry.

Prepare for AI-Accelerated Threat Activity

AI is lowering the barrier to tool creation and phishing sophistication.

- Train SOC teams to recognize AI-assisted phishing and deepfake techniques.
- Harden identity verification workflows beyond voice or SMS validation.
- Protect internal AI systems against prompt injection and data poisoning.

Defensive AI governance must mature alongside adversarial capability.

CHAPTER 7

Closing Perspective

The defining lesson of 2H 2025 is that attackers exploit trust - trust in identities, integrations, automation, and development ecosystems. They move quickly, leverage legitimate authentication flows, and monetize access at scale. Traditional perimeter defenses and reactive incident response models are no longer sufficient in isolation.

Organizations that treat identity governance, conditional access enforcement, and cloud-native visibility as foundational architecture, rather than supplemental controls, will be best positioned to reduce risk in 2026 and beyond. However, maintaining that level of vigilance requires sustained monitoring, rapid investigation, and continuous tuning of security controls across identity, endpoint, cloud, and SaaS environments.

For many organizations, this operational burden exceeds internal capacity. Modern Managed Extended Detection and Response (MXDR) services can help alleviate that pressure by providing continuous monitoring across identity and cloud control planes, rapid investigation of anomalous authentication activity, and proactive disruption of credential abuse before it escalates into ransomware or large-scale compromise. By combining platform telemetry, real-time threat intelligence, and experienced security analysts, an MXDR approach helps organizations close the gap between detection and response, particularly in identity-driven attack scenarios where minutes matter.

In an era where attackers log in rather than break in, continuous validation, rapid response, and expert oversight are no longer optional. They are essential to staying ahead of an adversary that is increasingly automated, opportunistic, and identity-focused.

Credits

To stay ahead of cyber threats, important trends and read our threat research, visit the Ontinue blog

Threat Research

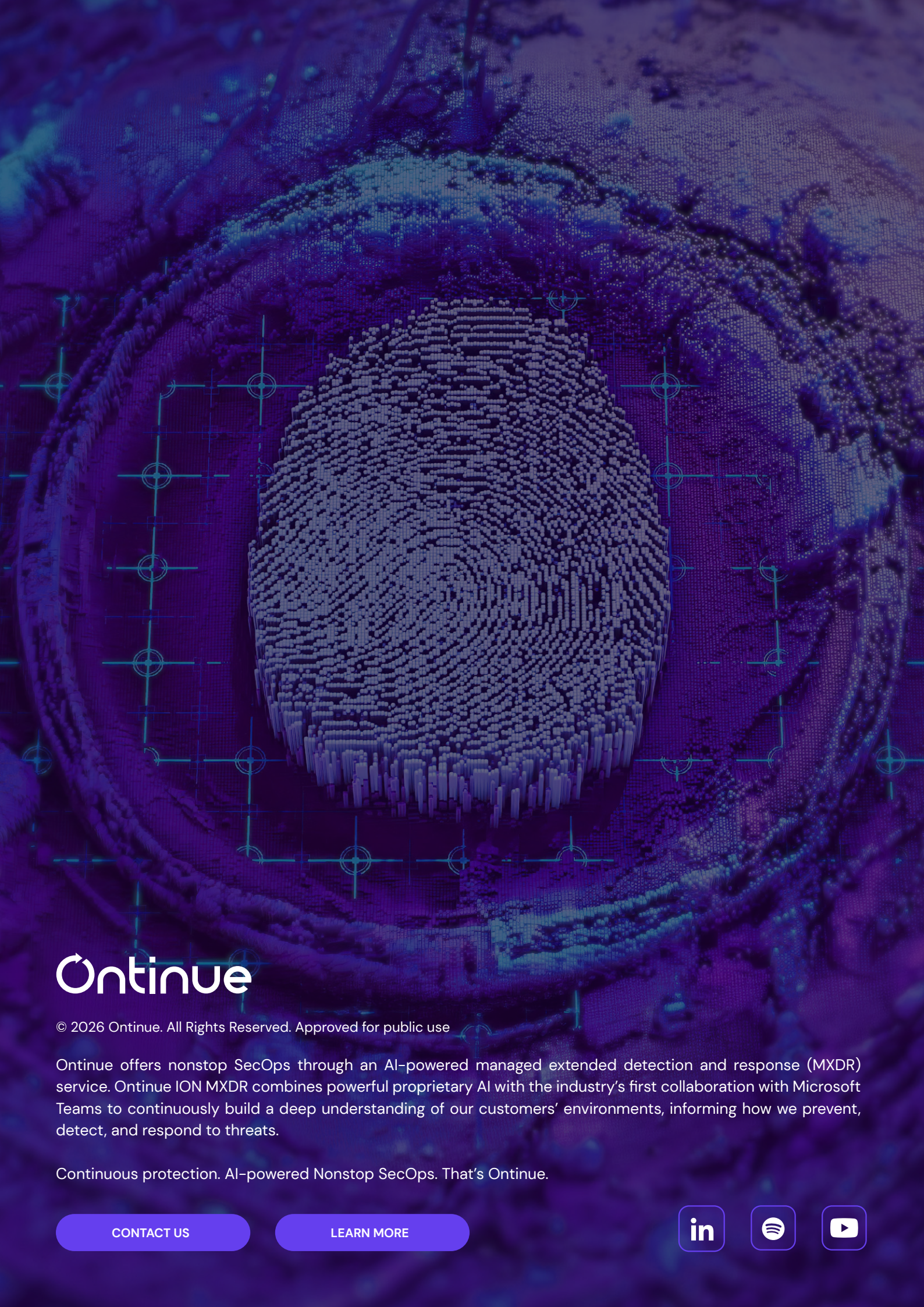
Rhys Downing
Domenico de Vitto
Balazs Greksza
Manupriya Sharma
Usha Sree Yannapu
Johnson Manual
Sharon NG
Ardit Hereqi
Kaan Varturk
Daley Oladapo
Paul Ducklin

Design and Editorial

Alison Raymond
Alberto Martinez
Alex Berger

Marketing and publishing

Advanced Threat Operations security research blogs



Ontinue

Ontinue offers nonstop SecOps through an AI-powered managed extended detection and response (MXDR) service. Ontinue ION MXDR combines powerful proprietary AI with the industry's first collaboration with Microsoft Teams to continuously build a deep understanding of our customers' environments, informing how we prevent, detect, and respond to threats.

Continuous protection. AI-powered Nonstop SecOps. That's Ontinue.

CONTACT US

LEARN MORE