MOUNT XION TECHNOLOGIES

Full report

ShinyHunters: How Access Became Data Theft

What the 2024-2026 campaigns reveal about human trust, integrations, and the chances to stop an attack.

Donte Wong, CISSP, Founder and CEO

September 23, 2026

Last updated October 6, 2026

Full report. A shorter article accompanies it at mountxiontechnologies.com.

Published: September 23, 2026. **Last updated:** October 6, 2026, with the update paragraph below, the Case 6 renewed-exploitation subsection, the three October timeline rows, and the FBIJobs update subsection. The original edition's evidence closed on September 22, 2026, 9:20 p.m. Central; the added material was checked on the morning of October 6.

Audience: Executive and technical leaders.

Scope: Selected public-source, defensive cases involving intrusions and data theft associated with, attributed to, or later branded as ShinyHunters activity from 2024 through 2026. No leaked records, personal data, exploit code, attacker infrastructure, or operational intrusion instructions are reproduced here.

The warning and the claim

On May 15, 2026, the FBI published a public warning about ShinyHunters. On September 22, a representative using that name told reporters it had broken into FBIJobs.gov, the bureau's hiring portal, and taken agents' and applicants' records. Reporters documented an apparent defacement and later an unavailable notice. The actor's stated demand was to retract or correct the May warning, according to Nextgov.

The FBI said it was investigating the claims. Its public statement did not confirm an entry path, the systems reached, or the volume taken. That distinction matters: a compelling story is not a completed investigation.

Update, October 6, 2026. The investigation has moved since this edition was first published on September 23. The FBI now says its review found that the incident resulted from a security failure on a platform managed by a third party, after a contractor did not apply a patch issued to secure it, and that it has removed the contractor. Police in the Netherlands and Jordan have detained suspects, and the FBI says it has worked with partners to arrest multiple subjects. The group has said it will not publish the data it claimed to hold. The analysis below is kept as it was first published on September 23. The update subsection in the FBIJobs case records the new evidence with the same labels, and what it means for the controls in this edition.

Across this study, stolen credentials, convincing calls, trusted integrations, privileged support workflows, excessive guest permissions, and exploitable software opened different paths to data. Some software attacks began before a patch existed. The human mistake was rarely the last chance to stop the theft. The defensive question is where the system could have interrupted what followed.

Executive assessment

This is a study of access becoming business harm, not a claim that one stable team carried out every event. Responders track separate operators and phases under identifiers including UNC5537, UNC6040, UNC6240, UNC6395, UNC6661, and UNC6671. The ShinyHunters name can connect an extortion or publication claim without establishing who obtained initial access.

The campaigns nevertheless expose a coherent defensive problem. Attackers repeatedly converted one weak point into trusted access:

- old credentials stolen from an unmanaged or previously infected endpoint;
- a convincing voice call that led a worker through a legitimate connected-app authorization process;
- a trusted third-party OAuth integration whose tokens reached many customers;
- untrusted customer content opened in a privileged support workflow;
- an internet-exposed enterprise application exploited before a patch existed; or
- excessive guest permissions on a public portal, without a stolen login.

After access, the attackers often used the victim's normal interfaces, sessions, tokens, APIs, export tools, and administrative functions. The theft looked less like destructive malware and more like an authorized user or application doing something abnormal at high speed. That makes human behavior important, but it does not make the employee the whole explanation.

The cases follow the attacker's choices through people, workflows, identities, applications, data, and business consequences. The recommendations are defensive design judgments. They are not findings that a named victim lacked every listed control.

Founder's Perspective: think like the attacker, then check your own assumptions

The weakest link is still human behavior, but that observation cannot become a way to transfer responsibility to an employee. The organization has to keep one mistake from becoming unrestricted access.

Before I founded Mount Xion, I handled this firsthand. I review the logs myself whenever phishing attempts are on the rise. That habit is how I caught it, within 90 minutes. In the age of AI, that is far too long. A man-in-the-middle phish had captured a user's sign-in. What stood out was impossible travel: the account signed in from the user's usual city, then from another state seconds later. Nothing acted on that signal. The organization's Entra sign-in policy blocked only locations outside the United States, so both sign-ins passed. That was the wrong posture. A country boundary is not a trust boundary.

The attacker got in and then went quiet. My read was that they were banking on no one noticing, so the credentials could be used later, on their schedule. The organization relied on a managed service provider for monitoring, but no system or policy enforced that kind of review. I ran the full trace anyway: what that identity had touched, what sessions and tokens it still held, and where else it could reach. There was no other foothold, but the session was still live. If I had not been doing that due diligence, the attacker could still be in that network today.

That review habit is what I am arguing for. The absence of activity is not the absence of an attacker. A password reset that leaves sessions and tokens alive is not a complete containment response. Assigning monitoring to a provider is not the same as knowing it happens. After the incident, monitoring became a defined, reported part of the provider's scope. The organization added automated alerts, so a sign-in like that one would be flagged without anyone having to look. It also tightened its sign-in controls. Until a system or policy enforces the review, someone who owns the risk has to look.

I came to civilian security from a military environment that spent more of its effort on the human and process side than industry does. That balance is worth importing. The technology community risks becoming so excited about new tools that it neglects the craft of thinking like a hacker at every level. That craft means following the attacker's choices through the person, the workflow, the identity, the application, the data, the operations, and the business consequence. Tools scale sound judgment. They do not replace it. Executives need enough adversarial understanding to govern the tradeoffs and recognize where a control can fail.

Donte Wong, CISSP, Founder and CEO, Mount Xion Technologies

Follow the access path

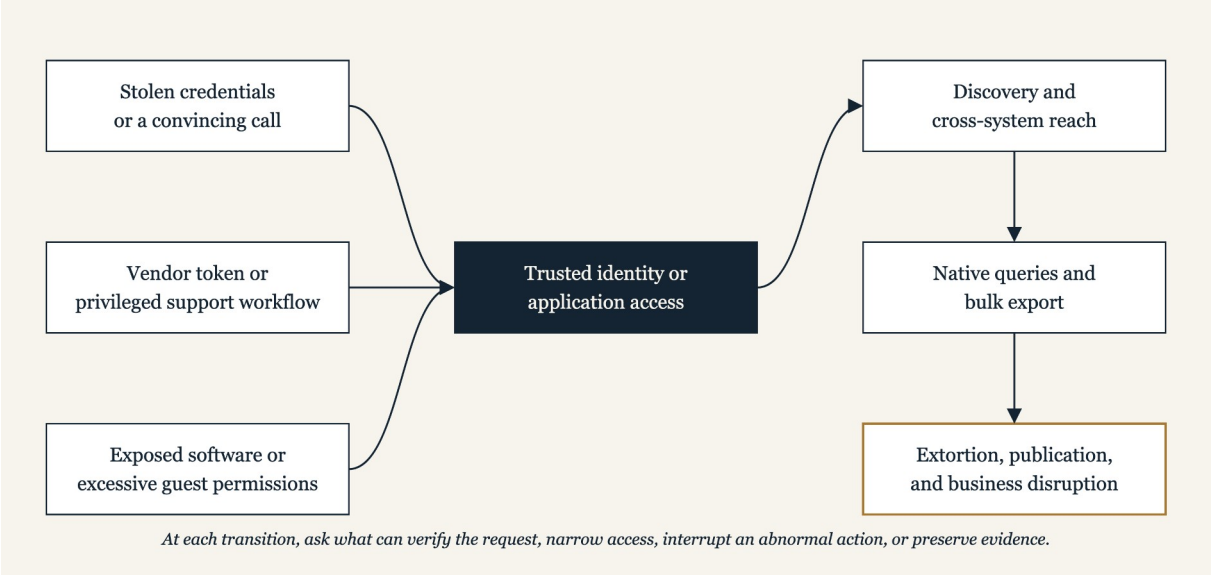


Figure 1. The access path across the 2024-2026 campaigns.

At each transition, ask what can verify the request, narrow access, interrupt an abnormal action, or preserve evidence. Guest access can expose records directly. Not every campaign requires a durable session or lateral movement.

How to read the evidence

This study separates incident facts from attacker statements and professional recommendations.

Label	Meaning in this study
Confirmed fact	Stated by the affected organization, a government authority, or a vendor or responder with direct incident visibility.
Corroborated observation	Independently observed or partially validated, but insufficient to prove source, scope, or attribution.
Direct attribution	An authoritative party ties a tracked actor or cluster to the activity through incident evidence.

Label	Meaning in this study
Brand claim or linked operation	ShinyHunters claimed, advertised, extorted, or published data, or another source linked the activity to that ecosystem. This does not prove end-to-end responsibility.
Campaign precedent	A behavior was demonstrated in an earlier, separately confirmed campaign. It is not proof that the same step occurred in a later incident.
Open theory	A plausible explanation that cannot be selected or rejected without additional telemetry.
Founder's Perspective	Donte Wong's professional assessment or recommendation. It is intentionally distinguished from a forensic finding about a named victim.

The study names organizations only when public evidence is strong enough to add defensive value. It does not reproduce leaked material, and it does not infer negligence from an incomplete public record.

Why attribution must be split by role

Mandiant and Google Threat Intelligence use different cluster identifiers when the evidence supports different operators, partnerships, or behaviors. Their January 2026 analysis specifically cautions that the ShinyHunters name can cover shifting alliances and possible impersonation.

Each case should therefore answer four separate questions:

Role	Question
Initial access	Who stole a credential, made the call, compromised the integration, or exploited the application?
Data theft	Who used the resulting access to discover and export data?
Extortion	Who contacted the victim or demanded money or another concession?
Publication	Who advertised, leaked, or claimed the data publicly?

Sometimes the public evidence answers all four. Often it does not. A shared name, forum account, or leak post is weaker evidence than victim telemetry, incident-response findings, or a criminal admission.

Timeline: evolution of the campaigns

Period	Event	Defensive lesson
February to October 2024	UNC5537 Snowflake customer-account campaign	Old credentials, password-only access, unrestricted locations, and unblocked bulk exports can turn an endpoint infection into a cloud data breach.

Period	Event	Defensive lesson
April to May 2024	AT&T, Ticketmaster, Santander, Advance Auto Parts, Neiman Marcus, and other customer environments affected	Centralized data copies and native export paths need least privilege, retention limits, and fail-closed anomaly controls.
March to September 2025	UNC6040 Salesforce vishing and connected-app campaign	Training must cover legitimate-screen and OAuth abuse, while app authorization and high-volume export require technical enforcement.
August 2025	UNC6395 Salesloft Drift token campaign	Non-human identities and vendor integrations require the same or greater control as privileged users.
November 2025 and June 2026	Gainsight and Klue integration incidents	Similar token-abuse mechanisms do not establish the same operator.
January 2026	UNC6661 and UNC6671 SaaS and SSO campaigns	A password reset is insufficient if sessions, tokens, OAuth grants, or new authenticators survive.
March 2026	Experience Cloud guest-access campaign	Review the data exposed to an anonymous visitor. No stolen login is required.
April to May 2026	Instructure Canvas incident	Systems must isolate untrusted content from privileged support sessions and treat near misses as signals for class-wide remediation.
May to June 2026	Oracle PeopleSoft CVE-2026-35273 campaign	Zero-day resilience depends on exposure reduction, segmentation, workload least privilege, egress control, and rapid containment.
July to August 2026 reporting	Microsoft's OAuth review and Google's UNC6671 update	Join identity and application telemetry while preserving source-specific actor attribution.
September 22, 2026	FBIJobs.gov claim	Communicate with evidence labels and investigate multiple hypotheses without turning campaign precedent into incident fact.
September 25, 2026	Renewed CVE-2026-35273 exploitation with a WAF bypass	A WAF rule is a stopgap. Apply the fix, and search the logs for the path and its encoded variants.
September 15 to October 5, 2026	Arrests in the Netherlands and Jordan; the FBI confirms multiple arrests	Arrests change who is willing to talk. They do not change the exposed component or its patch state.
October 5, 2026	The FBI attributes the FBIJobs incident to a patch not applied on a third-party-managed platform	A platform run by a contractor is still the organization's attack surface. Verify patch state; do not assume it.

PowerSchool is intentionally not treated as a ShinyHunters deep case. Its investigation report places the underlying intrusion in December 2024 and identifies compromised support credentials as the access path. The public evidence used for this study does not establish that a ShinyHunters operator performed that intrusion. The exclusion illustrates why a ransom note, forum identity, or leak-site name is not sufficient to attribute initial access.

Case 1: the 2024 Snowflake customer-account campaign

What happened

Mandiant tracked a broad campaign as UNC5537. Its June 2024 analysis found that the actor used credentials previously stolen by information-stealing malware to access customer Snowflake instances. It found no evidence that Snowflake's enterprise environment was breached.

Three conditions repeatedly enabled successful access:

1. stolen credentials remained valid, sometimes years after exposure;
2. affected accounts did not have MFA enabled; and
3. affected customer instances did not restrict access through network allow lists.

Mandiant reported that at least 79.7 percent of the accounts used by the actor had known prior credential exposure. Some infections involved contractor systems also used for personal activity. That is campaign-level evidence, not proof that an employee or contractor at any particular named victim behaved that way.

Once authenticated, UNC5537 needed neither malware nor a platform exploit. The actor signed in through Snowflake's own web interface and command-line client. It ran a reconnaissance utility Mandiant named FROSTBITE to list users, roles, and sessions, and used the DBeaver database client to query across instances. Then it followed the path any analyst would: SHOW TABLES to enumerate what existed, SELECT * on the tables it wanted, CREATE TEMPORARY STAGE to make a landing area, COPY INTO that stage with GZIP compression to shrink the files, and GET to pull them to its own machine. A temporary stage deletes itself when the session ends. None of that looks wrong to a database that has already accepted a valid password.

The scale is now supported by a criminal admission. In August 2026, the U.S. Department of Justice announced that Connor Riley Moucka pleaded guilty to participating in a scheme that compromised cloud-hosted data belonging to at least 165 SaaS customers, stole billions of records and terabytes of data, and extorted victims. DOJ said the conspirators received more than $2.5 million in ransom payments and caused victim-company losses exceeding $9.5 million, excluding losses to customers.

Representative impacts

AT&T. AT&T's SEC filing confirmed that files were exfiltrated from a third-party cloud workspace between April 14 and April 25, 2024. The files contained call and text interaction records for nearly all AT&T wireless customers and for customers of mobile virtual network operators using its network. The records covered May through October 2022, plus a single day in January 2023 for a small number of customers. The data did not include call or text content, Social Security numbers,

or dates of birth. AT&T later confirmed to TechCrunch that the workspace was hosted on Snowflake. This is a high-confidence campaign link, but AT&T did not publicly identify the specific person who obtained the credential or state that ShinyHunters performed every phase.

Ticketmaster and Live Nation. Live Nation reported unauthorized activity in a third-party cloud database primarily containing Ticketmaster data. A criminal actor offered alleged company user data for sale on May 27. Ticketmaster's incident notice said some North American customer information was affected. WIRED reported that ShinyHunters-branded accounts advertised the data, making the brand claim strong. The widely repeated figure of 560 million records came from the criminal offer and was not confirmed by Live Nation or Ticketmaster. The public evidence does not resolve whether the branded account performed initial access, partnered with the intruder, or sold the data later.

Santander. Santander's May 2024 statement confirmed access to a third-party-hosted database containing information about customers in Chile, Spain, and Uruguay, plus all current and some former group employees. It said the database did not contain transactional data, online banking credentials, or passwords. A ShinyHunters-branded actor advertised purported data, according to the same WIRED reporting, but Santander did not confirm the actor's record counts.

Advance Auto Parts and Neiman Marcus. The companies' breach notices identify unauthorized access to data in their Snowflake environments. Advance Auto Parts' California notice covers job applicants' Social Security, license, and birth-date information. Its SEC filing extends the affected population to current and former employees. Neiman Marcus disclosed names, contact information, and dates of birth: a narrower set, but still identity-relevant. These cases demonstrate that old human-resources and customer data remains valuable even when it is not part of a live production transaction.

Role-based attribution

Phase	Best-supported assessment
Initial access	UNC5537 used previously stolen customer credentials. Moucka admitted participation in the broader conspiracy.
Data theft	UNC5537 used native Snowflake functions to enumerate and export customer data.
Extortion and sale	DOJ confirms conspiracy-level extortion and advertising. ShinyHunters-branded identities claimed or sold data associated with some victims.
Publication	Brand and forum evidence varies by victim. It does not prove the brand holder performed initial access.

What could have interrupted the chain

- **Network policies on the warehouse.** Mandiant's summary is to limit traffic to trusted locations for crown-jewel data. A credential stolen from a contractor's laptop is worth much less if the instance only answers from the corporate egress or a VPN range.
- **No single-factor path, and no stale secret.** Every affected account authenticated with a password alone. Infostealer output stayed valid for years. Enforce MFA at the platform level,

check corporate identities against exposure feeds, invalidate on match, and investigate the endpoint that leaked the credential.

- **Alert on the unload sequence, not the login.** Enumeration, stage creation, a compressed COPY INTO, and a GET from a new address make up this campaign's signature. Query history records each command as it is submitted. A successful GET status means Snowflake authorized the transfer, not that the download completed. The client address lives in login history, so confirm a completed download against the session and login records. A hold on bulk unloads from unfamiliar clients, with evidence retained outside the warehouse, turns that record into a stop.

The credential exposure may have begun with human behavior, but later controls still mattered: invalidating exposed credentials, enforcing MFA, restricting network access, and limiting exports. The incident chain identifies several opportunities to intervene after the original theft.

Case 2: 2025 Salesforce vishing and connected-app authorization

What happened

Google Threat Intelligence tracked the callers and initial-access operators as UNC6040. Its June 2025 analysis described actors who impersonated IT staff by phone and walked employees to Salesforce's connected-app setup page. There the employee entered a connection code that linked an actor-controlled copy of Data Loader, sometimes renamed as a ticketing tool, to the company's environment. Data Loader is a legitimate mass import and export utility. It typically needs the user's profile to carry the "API Enabled" permission. The resulting OAuth grant enabled native API access and bulk export. This was not exploitation of a Salesforce software vulnerability.

In some incidents, the operators also used Okta-themed pages to obtain credentials and MFA responses, then moved into other services such as Okta or Microsoft 365. Follow-on extortion could begin weeks or months later. Google tracked that phase separately as UNC6240, which consistently claimed the ShinyHunters name. The FBI's September 2025 FLASH likewise described UNC6040 vishing, malicious app authorization, API queries, bulk exfiltration, and later ShinyHunters-branded extortion.

Google was also affected. In an August 5 update, it disclosed similar UNC6040 activity against a corporate Salesforce instance in June 2025. The instance held small-business contact information and notes. Google said the retrieved material was basic, largely public business data and that it cut off access after a short window.

Qantas as the human-workflow case

The Australian privacy regulator's Qantas inquiry found that a third-party contact-center agent received a call from someone impersonating Qantas IT support. The caller directed the agent to a legitimate CRM-related site and walked the agent through actions represented as necessary to close a support ticket. Those actions connected the CRM environment to an attacker-operated extraction tool.

Qantas detected unusual failed-login alerts and contained the access on June 30,

1. Approximately 5.67 million customer records worldwide were compromised.

The affected CRM did not store payment-card, financial, or passport data. Qantas reported that passwords, PINs, and login details were not accessed or compromised.

The regulator's report records preliminary inquiries, not a concluded investigation or final regulatory view. During that preliminary work, the OAIC did not identify a likely privacy-control violation. The report is still important because it records that Qantas already had recurring training, supplier audits, role-based access, and incident response. It also recognized that standard awareness training usually emphasizes credential theft, not being guided through a legitimate app-authorization interface. The operative weakness was that an end user could authorize a third-party application with meaningful access. The relevant platform default was later changed.

Qantas did not publicly name UNC6040 or ShinyHunters. The incident follows the same connected-app vishing pattern, but the responsible individuals and later extortion relationship should remain campaign-linked rather than victim-confirmed.

Role-based attribution

Phase	Best-supported assessment
Initial access	UNC6040 conducted the campaign's vishing and phishing operations. Qantas confirmed the human-workflow mechanism but did not name the actor.
Data theft	UNC6040 used connected-app or API access across the broader campaign. Qantas confirmed extraction through an attacker-operated tool.
Extortion	Google separately tracks the later extortion cluster as UNC6240.
ShinyHunters brand	UNC6240 consistently claimed the name. Public evidence does not establish that the callers, data thieves, extortionists, and publishers were always the same individuals.

What could have interrupted the chain

- **Take "API Enabled" off the profile.** Google and Salesforce both say to remove it from profiles and grant it through a permission set only to the people who run mass operations. Without it, the Data Loader the caller connected cannot export.
- **Make connected apps an administrator decision.** Restrict "Customize Application" and "Manage Connected Apps" to a few administrators and allowlist approved apps, so no user can self-authorize a new one from a phone call.
- **Bind logins and app tokens to known networks.** Login IP ranges on profiles and "Enforce IP restrictions" on each connected app deny or challenge an authorization from an unexpected address.
- **Block the large download in flight.** Transaction Security Policies can block a large data download as it happens, and Event Monitoring records who queried what through the API.

Correlate a new OAuth grant with discovery and export activity, and test how quickly the grant can be revoked.

- **Verify the caller out of band.** An unsolicited IT call gets a callback on a known number before anyone opens a setup page. Train for abuse of legitimate consent screens, not only fake login pages.

Case 3: 2025 Salesloft Drift OAuth-token campaign

What happened

The Salesloft Drift case shows that the human target can be replaced by a trusted non-human identity. Salesloft's public chronology, reproduced in reporting on its statement, says an actor gained unauthorized access to its GitHub account. The actor later accessed the Drift AWS environment and obtained OAuth tokens used for customer integrations. The public record does not establish the exact bridge between those environments, so a specific secret or repository should not be declared the cause.

Google Threat Intelligence tracked the customer-access activity as UNC6395. Beginning as early as August 8 and continuing through at least August 18, 2025, the actor used compromised Drift tokens to access numerous customer Salesforce instances. It systematically exported data and searched the stolen material for AWS access keys, passwords, and Snowflake-related access tokens. It deleted its query jobs afterward, but the logs survived. Compromised tokens for the Drift Email integration also enabled access to a very small number of Google Workspace accounts. Salesforce itself was not exploited.

Cloudflare as the machine-speed case

Cloudflare's incident report shows the sequence and the speed. The actor counted the Account, Contact, User, and Case objects and pulled the schema for Case. It ran a Salesforce Bulk API 2.0 job against the Cases object from 11:11:56 to 11:15:18 UTC on August 17. It then called the same API to delete the job it had just run. The traffic came from ordinary automation user agents, including one that named an open-source secret-scanning tool. Support-ticket text and contact information were exfiltrated. Attachments and files were not.

Cloudflare then built regex and entropy scanners, found 104 of its own API tokens in the exposed case text, and rotated them. It found no suspicious use of those tokens and no compromise of Cloudflare services or infrastructure, disconnected every third-party integration from Salesforce, and moved to weekly secret rotation. The episode demonstrates how support systems become secret repositories and how quickly a valid integration token can move data.

Attribution caution

The authoritative public attribution is UNC6395. ShinyHunters later claimed or was reported to have a role, but the public evidence is contradictory. BleepingComputer reported contact with a person claiming ShinyHunters affiliation. It also reported Google's response that it had not seen compelling evidence for the connection, and that person's statement that the campaign described by Google was not theirs. Later reporting alleged a relationship with ShinyHunters and Scattered Spider-

associated actors. Those are reported links, not an authoritative identification of UNC6395 as ShinyHunters.

This case belongs in this study because it intersects with the brand and campaign ecosystem, and because it exposes the same defensive problem. It must not be rewritten as a definitively attributed ShinyHunters intrusion.

The integration problem continued

Microsoft's July 2026 review describes a November 2025 campaign against Gainsight-published Salesforce apps and a separate June 2026 Klue incident involving Storm-3138. Both involved trusted connections or credentials used to access customer Salesforce data. Those examples extend the integration-risk pattern. They do not extend the UNC6395 identifier beyond the Drift campaign or establish a common operator.

What could have interrupted the chain

- **Scope the token to the job.** Google's first recommendation is to review connected-app scopes and avoid full. A chat integration does not need to read the Case object. Review who may create or expand an export-capable connection.
- **Pin the integration to the vendor's addresses.** "Enforce IP restrictions" on the connected app refuses a token presented from an unfamiliar network, whoever holds it.
- **Watch the query pattern.** Event Monitoring's UniqueQuery events log the SOQL a connected app runs. COUNT() sweeps followed by a Bulk API job on a sensitive object is the campaign in one line, visible before the job completes. Retain that evidence outside the compromised application's control.
- **Treat support cases as a secret store.** Scan case text with pattern and entropy tools, rotate on a schedule, and treat a pasted key in a ticket as an incident. Google's recommended search of Salesforce objects covers AWS key identifiers beginning AKIA, Snowflake hostnames, the strings password, secret, and key, and the organization's own VPN and SSO login URLs.
- **Test the kill switch before the incident.** Know which administrator revokes a vendor's connected app across the enterprise and how long it takes.

The Microsoft 365 equivalent: app consent

The Microsoft 365 version of the same decision is app consent. Microsoft's recommendation is to let users consent only to apps from verified publishers and only for permissions classified as low impact, with the admin consent workflow routing everything else to an administrator. That setting does not touch existing grants, so the inventory of what has already been consented to is a separate job. Defender for Cloud Apps app governance covers that inventory across Microsoft 365, Google Workspace, and Salesforce. A policy there can alert on an app whose permissions or data use match a condition and disable it automatically. These are governance and investigation capabilities. None of them blocks every abnormal export on its own.

Founder's Perspective: near misses and access requests are both telemetry

Annual training completion is not the same as human-risk management. If staff keep receiving phish or vishing calls, ask what stopped each attempt and what pattern is developing. Ask what the business impact will be when one succeeds, and how to interrupt that path before it becomes an incident. Near misses are not empty events. They are attack telemetry.

Access requests deserve the same treatment. In a previous role, I approved a large IT vendor's request for full access to the organization's tenant. It was a routine ask from a trusted partner. Then it dawned on me that the organization had no need for the vendor to hold that level of access. The risk against the reward did not add up, so I rescinded the authorization. With third-party and supply-chain compromise growing, that should be the default: unless a concrete business need exists for a specific level of access, reduce it. A vendor integration starts with a decision to grant access. That decision needs an owner, a defined purpose, and a review when the risk changes.

Case 4: the January 2026 SSO, MFA, and SaaS expansion

What happened

In incidents spanning early to mid-January 2026, Google Threat Intelligence reported, UNC6661 operators called employees while victim-branded phishing pages synchronized with the caller's script. The attackers captured SSO credentials and MFA responses, registered an attacker-controlled device or factor, and used the resulting identity access to enter whichever SaaS applications the account could reach. Observed targets included environments associated with Microsoft 365 and SharePoint, Salesforce, DocuSign, and some Okta customer accounts.

The operators searched for sensitive documents and personal data using native SaaS functions. In at least one case, they used a Google Workspace add-on to delete the legitimate warning that a new security method had been enrolled. Compromised mail access could then support follow-on phishing.

Google maintained separate cluster identities. It linked later UNC6240 ShinyHunters-branded extortion to some UNC6661 intrusions through overlapping negotiation and stolen-data evidence. UNC6671 used similar vishing and SaaS access methods but showed different extortion behavior. This is evidence of a partnered ecosystem, not proof of one continuous operator.

Okta's threat research explained why number matching and one-time codes can still fail against a real-time caller and synchronized phishing kit. They reduce accidental approval but are not phishing-resistant authentication.

Google's August 6 update describes continued UNC6671 activity after BlackFile's alleged May retirement, including July targeting of financial and legal organizations. It reports calls to personal phones and lookalike SSO or passkey portals used to intercept credentials and MFA responses. Google assessed connections among several extortion brands while retaining competing explanations for that structure. This report keeps the activity under UNC6671. It does not establish that the operators are identical to the other ShinyHunters-linked clusters in this study.

What phishing resistant means in Microsoft 365

In a Microsoft 365 environment, “phishing resistant” has a precise meaning. Entra Conditional Access authentication strengths can require the built-in Phishing-resistant MFA strength, which accepts only Windows Hello for Business or a platform credential, a FIDO2 security key or passkey, or multifactor certificate-based authentication. A password plus a push, code, or text satisfies the ordinary MFA strength and not this one, which is exactly the gap a synchronized phishing kit exploits. For supported native applications, Conditional Access token protection then binds sign-in session tokens to the device that obtained them, so a token captured by a proxy cannot be replayed from another machine. Containment has to reach the tokens too. Microsoft’s own compromised-account guidance notes that app passwords survive a password reset. A separate session revocation invalidates refresh tokens and browser cookies. Continuous access evaluation carries that revocation to Exchange, SharePoint, and Teams in near real time. Reset the password, revoke the sessions, remove the unfamiliar authenticator, and only then call it contained.

What could have interrupted the chain

- **Require an authentication strength, not just MFA.** Apply a Conditional Access policy with the Phishing-resistant MFA strength to sensitive access and to enrollment and recovery paths, on managed devices. That strength admits only Windows Hello for Business or a platform credential, a FIDO2 key or passkey, or multifactor certificate-based authentication.
- **Require a managed device or a compliant network at sign-in.** A Conditional Access policy that requires a compliant or hybrid-joined device is evaluated before a token is issued, so a phishing proxy or a rented server with no device credential is refused wherever it sits. The Global Secure Access compliant network check closes the same gap for the network with no address list to maintain. Microsoft documents that a stolen session token replayed from outside that network is denied immediately. Country and address block lists are the weaker gate, because an attacker can rent infrastructure inside the allowed region.
- **Verify recovery independently.** High-risk resets and new authenticator enrollment get out-of-band verification. A risk policy forces phishing-resistant reauthentication at medium sign-in risk.
- **Contain the identity, not the password.** Revoke sign-in sessions and refresh tokens, delete app passwords, remove unauthorized factors and OAuth grants, and correlate the account recovery with subsequent SaaS access.

Additional case: a public portal with excessive guest access

Salesforce’s advisory, published March 7, 2026 and updated March 11, describes a campaign against overly permissive Experience Cloud guest profiles. Anonymous visitors queried CRM objects that customers had not intended to expose, without logging in, using a modified version of Aura Inspector, an open-source audit tool, at scale. Salesforce attributed the weakness to customer-configured guest settings, not a platform vulnerability.

FINRA’s alert names ShinyHunters. Salesforce’s public advisory does not name the group. The mechanism is established by the vendor. The actor naming here is attributed to FINRA. This is a

sixth access pattern: the application's guest permissions provide access without first deceiving a user or stealing a login.

What could have interrupted the chain

- **Audit the guest profile as if it were an employee.** Review the guest user profile's object and field permissions, turn on "Secure guest user record access," and set org-wide defaults to Private for guests.
- **Close the API side.** Disable "Allow guest users to access public APIs," uncheck "API Enabled" on the guest profile, and disable self-registration where it is not needed, testing required portal functions before each change.
- **Inspect what an anonymous visitor can actually retrieve.** Review the Aura event logs for anomalies and repeat the inspection on a schedule, not only after an advisory.

Case 5: the 2026 Instructure Canvas incident

What happened

The Canvas incident connects software weakness, normal employee behavior, token design, and cross-tenant privilege. Instructure's incident hub and customer FAQ describe two related attacks.

In the first, the actor created a Free-for-Teacher account and submitted a support ticket containing malicious content. When a customer-service agent opened the ticket, a cross-site scripting flaw executed in the agent's session, allowing the actor to obtain an authorization token with elevated, cross-customer access. In the second event on May 7, the actor used another then-unpatched XSS weakness in a discussion feature and an OAuth flow to generate a new token.

Enhanced monitoring detected and disabled the second attack in approximately ten minutes. Instructure says it found no data exfiltration during that second event. The actor altered customer pages and custom styling across roughly 300 organizations and disrupted some identity-provider configurations. Confirmed exposed data categories included usernames, email addresses, course names, enrollment information, and messages. Instructure reported no evidence that passwords, dates of birth, government identifiers, financial information, core course content, or submissions were exposed, and no evidence of lateral movement outside Canvas.

Instructure says the incidents appear to have been perpetrated by ShinyHunters. The FBI's May 15 public service announcement described the group's extortion pattern and warned that such activity can combine real, exaggerated, or fabricated claims with harassment and pressure. It cites no record counts or data volumes. Actor publicity does not establish incident scope. The affected provider's disclosures above supply the stated scope for this case.

Why this is not simply an employee mistake

Opening a customer support ticket was normal work. A software vulnerability caused untrusted content to execute in a high-trust support context. The resulting token then carried excessive reach. Training an agent not to view a ticket is not a credible primary control for a support platform.

The second event is equally instructive. Monitoring recognized the renewed behavior and contained it in about ten minutes, apparently preventing additional exfiltration. That is the practical value of analyzing behavior and turning a signal into an immediate response.

What could have interrupted the chain

- Isolate untrusted support content from privileged sessions and test the broader XSS class after finding one vulnerable path.
- Limit support-token reach across customers and require additional checks for unusually broad administrative actions.
- Monitor abnormal support-account behavior and exercise rapid containment, using the second incident as evidence of the value of intervention.

Human behavior analytics should be treated with the same seriousness as network analytics. A support session that begins behaving outside its purpose is a signal to investigate and contain, even when opening the original ticket was part of an employee's job.

Case 6: the 2026 PeopleSoft zero-day campaign

What happened

Oracle's June 10 security alert described CVE-2026-35273, an unauthenticated remote-code-execution vulnerability affecting supported PeopleSoft PeopleTools 8.61 and 8.62. Oracle assigned a CVSS score of 9.8 and urged immediate mitigation or patching.

Google Threat Intelligence had already observed UNC6240, the extortion cluster it tracks under the ShinyHunters name with the caveats above, exploiting the weakness from May 27 through June 9. Its campaign analysis described activity consistent with exploitation of the Environment Management Hub, followed in some incidents by remote-management deployment, environment reconnaissance, internal mapping, attempts to use discovered credentials, data staging, compression, defacement, and publication.

The exposed component was the Environment Management Hub, reached through the /PSEMHUB/hub path, alongside the Integration Broker listening connector at /PSIGW/HttpListeningConnector. Google's report states that both are administrative or system-to-system components and that blocking them at the perimeter is non-breaking for the browser sessions end users actually use. After exploitation the actor deployed MeshCentral remote-management agents disguised as cloud endpoints. It read the PeopleSoft process scheduler and WebLogic configuration files to map the environment. It sprayed a hard-coded list of administrative credentials over SSH at internal hosts named in the local hosts file. It also staged and compressed the data and dropped a defacement marker.

Google notified more than 100 organizations it assessed as potentially exposed, approximately 68 percent of them in higher education. Potential exposure is not the same as 100 confirmed victims. CISA added the vulnerability to its Known Exploited Vulnerabilities catalog on June 12 and set a June 15 federal remediation deadline.

NAIC as a confirmed impact case

The National Association of Insurance Commissioners reported that it detected unauthorized PeopleSoft access on June 11. It said the intruder obtained information that enabled temporary access to certain data-storage areas and later published data. Confirmed scope included public statutory financial reports, credit-rating information, and potentially routine technical information. NAIC said it had no current evidence that personal information, payment-account data, or confidential insurer or company information was exposed. It disputed the amount and scope of data claimed by the group. The provider's statement does not verify the actor's advertised volume.

Operational consequences included pausing rating-data feeds and temporarily suspending some designations. This shows that confidentiality is not the only measure of impact. Containment can interrupt authoritative business processes even when the confirmed data is less sensitive than the actor claims.

What could have interrupted the chain

The first wave began before a public patch existed. Controls beyond patch speed were therefore essential:

- **Take the administrative endpoints off the internet.** Disable the Environment Management Hub where the configuration allows, or block /PSEMHUB/* and /PSIGW/HttpListeningConnector from external networks at the firewall. Audit the WebLogic access logs for external POST requests to those paths and their percent-encoded variants, and watch for outbound SMB from PeopleSoft hosts.
- **Segment the application from the administrative and storage planes.** Constrain the workload identity. Deny outbound connections by default so a staged archive has nowhere to go. Keep the hosts file and configuration from doubling as a map for lateral movement.
- **Apply the mitigation, then hunt.** Once Oracle's remediation exists, apply it promptly. Then look for earlier access, remote-management agents, persistence, and abnormal transfers, because a patch does not remove what was established before it.

Google cautioned against relying on WAF request-body inspection alone. Patching and a retrospective investigation answer different questions.

Patching is direct threat reduction. Once remediation exists, an exposed, unremediated system can give an attacker a known entry path. Teams must also hunt for earlier access and limit what the application can reach. A patch does not remove persistence already established before it was applied.

Renewed exploitation after the patch

On September 25, 2026, Google reported renewed mass exploitation of CVE-2026-35273 by UNC6240. Many WAF rules written after the June alert matched the literal hub path before decoding it, while the application server decoded the request and routed it normally, so a request with one percent-encoded character reached the vulnerable component. Google's incident responders found JSP web shells on dozens of systems across higher education, technology, IT services, healthcare, agriculture, transportation, and government. Google's recommendations: apply Oracle's update; disable the hub or remove its application where the configuration allows;

search the WebLogic access logs for the hub path and its percent-encoded variants; inspect the hub's application directory for unauthorized files; rotate credentials reachable from PeopleSoft service accounts; and check for remote-management agents. Google's wording: WAF rules and path-based blocking are not a substitute for patching.

The renewed campaign turns Case 6 from a zero-day story into an n-day story. The first wave tested exposure reduction and segmentation. The second tested whether organizations had applied the fix or had only placed a rule in front of the flaw.

Developing case: the claimed FBIJobs.gov compromise

The FBIJobs event is the unresolved end of this study, not a confirmed ShinyHunters success story. It tests whether the evidence discipline applied to earlier cases will hold when the facts are incomplete and the claim is designed to create pressure. An update dated October 6, 2026 follows the open theories; the September 22 record stands as written.

What is known, claimed, and unknown

Assertion	Public evidence as of the cutoff	Assessment
FBIJobs.gov was the subject of unauthorized-activity claims.	The FBI told reporters it was aware of claims affecting FBIJobs.gov and was investigating.	**Confirmed**, but only as an investigation of claims.
The jobs portal was disrupted and displayed attacker-branded content.	Nextgov reported maintenance messaging and an earlier apparent seizure notice.	**Corroborated observation.** The exact cause and level of access remain unconfirmed.
A sample contained information associated with FBI or DOJ personnel.	Reporters checked portions of an approximately 5,000-record sample against public or previously compromised information.	**Partially corroborated.** Provenance, recency, completeness, and source remain unknown.
A new PeopleSoft zero-day enabled code execution.	A ShinyHunters representative described an undisclosed flaw to reporters. No public Oracle advisory tied a new flaw to this event by the cutoff.	**Actor claim.** Do not label it CVE-2026-35273 without evidence.
The entry point was the applicant portal's PeopleSoft Environment Management Hub.	The group gave CyberInsider a screenshot of a page under the /PSEMHUB/ path on apply.fbijobs.gov that appears to show Linux system information, and described that system as its entry point.	**Actor claim with an actor-supplied screenshot.** Unverified. If authenticated, it could support access involving the hub the June campaign exploited (Case 6). It would not establish the initial-access route or which flaw, if any, was used.
The same flaw is now being used against other organizations.	BleepingComputer relayed the group's claim that it is now using the alleged flaw against other organizations, including Fortune 500 companies. CyberInsider relayed a stated plan to use it against a broader range of targets.	**Actor claim.** No advisory, CVE, or affected organization confirmed it by the cutoff. The Case 6 controls apply either way.

Assertion	Public evidence as of the cutoff	Assessment
The actor tried to erase evidence from compromised servers.	BleepingComputer relayed the group's claim that it tried to erase evidence to make the flaw harder to identify.	**Actor claim.** It is why log preservation outside the affected systems is on the evidence list below.
The known June CVE-2026-35273 may have been reused against an unremediated system.	POLITICO reported that investigators were considering whether the earlier exploit was reused against a system that had not received Oracle's remediation or whether another path was used.	**Open theory.** Current PeopleTools version, remediation state, exposure, and initial access remain unconfirmed.
Investigators reportedly find the claims credible.	POLITICO cited two people with knowledge of the breach who said investigators believe the group's claims are credible.	**Unnamed-source reporting.** Not an official finding. It raises the weight of the claim without confirming scope or path.
The actor pivoted into FBI-managed AWS GovCloud systems.	Nextgov relayed the actor's account of movement into GovCloud.	**Actor claim.** No public cloud, identity, or network telemetry verifies the path.
Two to three terabytes were stolen, covering employees and applicants.	Nextgov attributed the volume to the actor and did not verify the entire dataset.	**Actor claim.** A small sample cannot establish the total.
Other FBI services were compromised.	Nextgov reported actor claims concerning human-resources systems and Medlink. BleepingComputer and CyberInsider relayed a claimed "Criminal Justice" service as well.	**Actor claim.** No affected-system owner confirmed that scope by the cutoff.
The aim was to force correction or removal of the May warning.	Nextgov reported the actor's stated demand and its one-week deadline.	**Actor claim about motive.** A demand does not establish the full motive or validate the claimed intrusion.

The source search recorded for the 9:20 p.m. cutoff found no substantive official update beyond the FBI's acknowledgment of its investigation, whether from the FBI, DOJ, Oracle, AWS, CISA, or an affected-person notice. The additional reporting expanded the plausible entry paths but did not confirm one. That means no official update was found, not that none exists. A pre-publication check on the morning of September 23 (the FBI and DOJ newsrooms, IC3 announcements, CISA news, Oracle security alerts, and AWS security bulletins) again found none. The FBI's May 15 warning remained posted.

What historical FBI records establish

The FBI's 2022 privacy impact assessment described FBIJobs and Candidate Gateway as internet-accessible components hosted on AWS GovCloud virtual machines and using Oracle PeopleSoft HRS 9.2, Oracle Database 19c, and Drupal. It described an accredited cross-domain connection through which completed applicant information moved to the FBI's HR Source system. A later FBIJobs FAQ described apply.fbijobs.gov as an Oracle application portal connected to, but distinct from, the public FBIJobs site.

These records make a PeopleSoft-centered incident plausible. They do not prove the September 2026 architecture, configuration, data retention, PeopleTools version, or attack path.

Three distinctions are critical:

1. PeopleSoft HRS 9.2 is an application release. It does not establish that the underlying PeopleTools version was 8.61 or 8.62.
2. Candidate Gateway was already described as GovCloud-hosted in 2022. Access to an FBI-managed workload inside GovCloud would not, by itself, prove compromise of AWS infrastructure or the AWS control plane.
3. A historical transfer path into HR Source does not prove that the actor crossed it, reached FBINet, or accessed the full employee population.

The open theories

Three broad initial-access theories fit the limited evidence:

Theory A: reuse of the known June vulnerability. CVE-2026-35273 was exploited as a zero-day before Oracle's June 10 alert and remediation. It could have been reused if an affected, exposed PeopleTools component remained unremediated. POLITICO reported that investigators were considering this possibility. The historical HRS 9.2 reference does not establish the PeopleTools version or whether the affected component was present, exposed, or patched.

Theory B: a different public-facing application flaw. The actor claims it used a new, previously undisclosed PeopleSoft zero-day. The earlier campaign makes the capability plausible, but no new Oracle advisory, CVE, or forensic evidence had confirmed that path by the cutoff.

Theory C: identity or administrative compromise. The actor obtained a valid credential, session, token, or privileged support path and used legitimate administrative functions to alter the site or reach data. ShinyHunters-linked identity campaigns make this plausible, but no public evidence selects it.

The actor-supplied screenshot does not choose among these theories. The group calls the Environment Management Hub, the component exploited in the June campaign, its entry point. If authenticated, the screenshot could support access involving the hub. It would not establish the initial-access route or which flaw, if any, was used. It fits a reused June flaw, a new flaw in the same component, and access that began elsewhere and reached the hub later. For defenders the consequence is the same under every theory: the Case 6 hub controls apply now, whatever the investigation finds.

A separate possibility is that the sample combines data from more than one source or time period. Matching public information proves neither fresh access nor claimed provenance.

The following evidence is needed before the case can be closed:

- current PeopleTools and application versions;
- the initial-access vector and first affected identity or host;
- whether access remained in Candidate Gateway or reached HR Source or other systems;
- whether any AWS service was compromised, as distinct from an FBI-managed workload hosted on AWS;
- verified exfiltration scope, timestamps, source systems, and record recency;
- whether logging was disabled, deleted, or preserved in an independent plane;

- an Oracle advisory or CVE for the claimed new flaw; and
- an FBI or DOJ incident update, affected-person notice, or independent forensic report.

Until that evidence exists, the current event should be described as a serious, developing claim with limited corroboration.

Update, October 6, 2026: what changed after the cutoff

The September 22 table, theories, and evidence list above are kept as written. This table records the public evidence added between the cutoff and October 6, 2026, 9:45 a.m. Central, with the same labels.

Assertion	Public evidence as of October 6	Assessment
The incident resulted from a patch not applied on a third-party-managed platform, and the contractor was removed.	Brett Leatherman, assistant director of the FBI's Cyber Division, in a statement reported by Reuters on October 5 and carried by SecurityWeek: the review had so far determined that the incident "occurred as the result of a security failure of a platform managed by a third-party organization" after "a contractor failed to implement a security patch explicitly issued to secure the platform," and the FBI "has removed the contractor."	**Confirmed as the FBI's current finding.** The FBI named neither the platform nor the contractor, and it has not named the flaw. "To date" leaves room for revision.
The platform was Oracle PeopleSoft.	Reuters' two unnamed sources, as relayed by SecurityWeek and The Hacker News.	**Reported, anonymously sourced.** Consistent with the 2022 privacy impact assessment and the actor's screenshot; not an FBI statement. The same sources named the contractor. This edition omits the name because the FBI did not give it, and the edition's rule is defensive judgments, not findings against a named organization.
Employee data was taken, and employees were warned.	The FBI told reporters on September 29 that it was in regular communication with affected individuals (CBS News). The Record reported an internal memo warning employees of the breach and the potential danger to them and their families.	**Confirmed in part.** The FBI acknowledges affected individuals; it has published no count and no data categories. Descriptions of what the records contain remain reporters' characterizations of actor-supplied samples.

Assertion	Public evidence as of October 6	Assessment
A suspect was arrested in the Netherlands.	Dutch police announced on September 29 the September 15 arrest of a 24-year-old Amsterdam man suspected of taking part in the group; a court ordered at least 90 days of pretrial detention (CBS News, citing the police statement). The FBI described him as "one of the alleged leaders."	**Confirmed.** Police did not name him, and this edition does not.
The group breached more than 140 organizations and took at least $70 million in extortion payments.	Leatherman's statement of September 29: "this cybercriminal and his co-conspirators have allegedly breached more than 140 organizations" and taken "at least $70 million in extortion payments" since 2025 (CBS News; Nextgov).	**FBI allegation.** Investigative figures, not adjudicated findings.
A second suspect was detained in Jordan and is cooperating.	Reuters, October 3, citing three people familiar with the matter; Jordanian state media confirmed an arrest without naming the suspect (The National).	**Confirmed arrest; reported cooperation.** The name and the cooperation come from unnamed sources, so this edition carries neither.
The FBI has arrested multiple subjects.	FBI statement to Reuters on October 3, repeated to The Register on October 5: the bureau "continues to aggressively investigate the recent cyber incident allegedly involving ShinyHunters, having already worked with partners to arrest multiple subjects."	**Confirmed.** Note the FBI's own "allegedly involving ShinyHunters."
The actor resumed exploiting CVE-2026-35273 with a WAF bypass.	Google, September 25 (Case 6).	**Confirmed by a responder with incident visibility.**
The actor used that bypass against FBIJobs and also a second, undisclosed flaw.	The group's statement to BleepingComputer, September 26.	**Actor claim.** The FBI's account fits a known, unapplied patch. No advisory supports a second flaw.
The group will not publish the FBI data.	The group's statements of September 28 and 29 (CyberInsider; Nextgov); the FBI entry was removed from its leak site.	**Actor statement.** No release had been reported by the cutoff. A withdrawn threat does not establish that the data was deleted.
The May 15 warning was retracted or corrected.	The IC3 page was checked on October 6 and carries no revision.	**Not done.** The actor's stated demand was not met.

Theories, as of October 6. The FBI's account matches Theory A in substance: a fix existed and was not applied. The FBI has not named the flaw. Google's report of the same actor's renewed CVE-2026-35273 exploitation and the actor's own statement make that vulnerability the leading candidate, and nothing official has named another. Theory B survives only as an actor claim with no advisory. Theory C is not selected by any evidence.

Evidence list, as of October 6. The initial-access vector is answered in kind (an unapplied patch on a third-party-managed platform) but not in detail. The FBI has acknowledged affected

individuals without a count. The Oracle advisory for a claimed new flaw remains absent. The remaining items (current versions, reach beyond Candidate Gateway, any AWS service compromise, exfiltration scope and recency, log preservation) remain open.

What it means for the controls in this edition. A platform run by a contractor is still the organization's attack surface, so patch state must be verified rather than assumed, and the contract must make that verification possible. A WAF rule is a stopgap, not remediation. Exposure reduction for the administrative hub holds under every theory. The arrests change who is willing to talk; they do not change the exposed component.

Cross-case finding 1: human behavior is security telemetry

Traditional programs often count annual course completion, phishing clicks, and reported messages. Those measures can be useful, but they are too narrow. The 2024-2026 cases show that attackers target workflows, not just inboxes:

- contractors use devices outside the organization's normal management plane;
- callers guide employees through real interfaces;
- help desks reset passwords or replace MFA methods;
- users grant OAuth access to legitimate-looking applications;
- support agents open content because their job requires it;
- privileged users create keys, integrations, or export jobs; and
- service identities behave like trusted software until a token is stolen.

A human-risk analytics program should join identity, device, help-desk, email, SaaS, data, and incident signals. It should ask:

1. What stopped the last attempt?
2. Which roles, teams, shifts, suppliers, or workflows are repeatedly targeted?
3. Are attempts becoming more persuasive, faster, or better synchronized?
4. What access would the targeted person or process create if the next attempt succeeded?
5. Which control can interrupt the path before success rather than merely report it afterward?

Useful measures include suspicious help-desk requests, denied MFA changes, aborted OAuth grants, repeated vishing by business role, reporting speed, high-risk actions after password reset, new-factor enrollment, anomalous token use, and export attempts stopped before data left.

Training should become shorter, recurring, role-specific, and tied to observed behavior. Realistic vishing, account-recovery, OAuth-consent, and vendor-support exercises should test both employees and the security team. When an employee is deceived, immediate non-punitive coaching creates more learning and more reporting than a culture built around blame.

Cross-case finding 2: risk-adaptive controls should fail closed

The product-neutral model is **risk-adaptive access control driven by behavioral analytics**. In Microsoft environments, Entra ID Protection and risk-based Conditional Access are examples of this pattern. They can combine user or sign-in risk with access decisions such as blocking, requiring stronger authentication, or requiring remediation. The larger design applies across identity, SaaS, cloud, and data platforms.

Microsoft's July 13 analysis recommends connecting Salesforce to Defender for Cloud Apps for visibility and threat detection. With Salesforce Shield Event Monitoring, its upgraded connector supplies additional activity context. Microsoft also describes connected-app permission review, inactive-app discovery, and app risk scores. These are investigation and governance capabilities. This study does not assert that enabling the connector automatically blocks every abnormal export.

Abnormal event	Medium-risk response	High-risk response
Authentication from an unmanaged device or suspicious location	Require phishing-resistant reauthentication and a managed device.	Block, revoke the session, and open an incident.
MFA replacement immediately after password reset	Pause the enrollment and require independent verification.	Block the new factor, revoke sessions and tokens, and investigate account recovery.
Unusually large download or API export	Pause the export and require an independent approver to release it.	Terminate the export, revoke the token or session, preserve evidence, and open an incident.
Creation of an access key, token, service credential, or high-impact OAuth grant	Time-limit it, restrict scope, and require independent approval.	Deny creation, contain the initiating identity, and investigate.
Attempt to disable logging or a core security control	Not applicable. Security-control tampering should not be treated as an ordinary medium-risk event.	Always block, alert through an independent channel, preserve immutable evidence, and require incident handling.

Risk thresholds must be tuned to the business. A hospital, financial institution, small manufacturer, school, and global cloud provider will not make identical availability tradeoffs. The decision should be explicit, tested, and owned, not left as the accidental default of an alerting product.

Cross-case finding 3: two-person integrity for business-state changes

The four-eyes principle requires two independently authenticated and authorized people for changes that can erase evidence, remove safeguards, or alter the organization's trust boundary. It should cover at least:

- deleting audit logs or reducing retention;
- disabling foundational security controls;
- changing federation, tenant ownership, or root trust;

- mass-deleting business data;
- creating or materially expanding high-impact integrations or credentials; and
- modifying break-glass accounts.

A normal administrator modifying a break-glass account should require dual approval. A protected break-glass identity may need to repair another emergency account during a declared incident. It should do so only through a narrowly defined process: time-limited privilege, immediate off-path alerting, immutable records, and independent review as soon as another administrator is available. Compromise of one emergency account must not allow silent takeover or destruction of every other emergency path.

Microsoft Entra Privileged Identity Management can require an approver for role activation. Microsoft's role-settings guidance recommends having at least two available approvers. That is an availability recommendation, not native two-person approval: the first approver can resolve a request. True two-person integrity therefore requires a separate workflow or enforcement mechanism.

The policy must reflect actual staffing, operating hours, and emergency needs. A dual-control rule that blocks the only available responder can increase risk and will eventually be bypassed. Organizations should define alternates, escalation timers, narrow emergency authority, and retrospective review before the crisis.

Security controls enforce business governance. They do not replace it. Where law, regulation, litigation hold, or contract requires preservation, mandatory retention locks can supersede executive preference. Outside those constraints, collective leadership may authorize exceptional deletion through a documented process involving legal, privacy, compliance, and security review, while two administrators execute the approved action. Executive authorization establishes internal authority. It does not make unlawful destruction lawful.

Cross-case finding 4: bounded autonomous defense with a human on the loop

Attackers do not wait for a change board. A defensive system loses much of its value if every reversible containment step must wait for a person while an attacker exports data in minutes.

Google's September 8 report describes UNC6240 using Claude Code for intrusion-related coding and custom Model Context Protocol (MCP) tools to analyze stolen directories for extortion. The table rows name Claude Code, while the caption names Gemini. That labeling inconsistency in the source remains unresolved.

Anthropic's September report describes separate operators it assessed as suspected ShinyHunters affiliates. For one unnamed supply-chain operator, it reports a dump containing over 2,100 Azure AD token sets across more than 40 corporate tenants in about 34 hours, with AI agents performing nearly all of that work. It also reports a separate compromise by that operator reaching cloud administrative control from one developer token in roughly three hours. The report does not give the same AI work-share estimate for that second incident.

These are source-attributed operations, not proof that every ShinyHunters-linked intrusion used the same tools or autonomy. Neither passage establishes AI use in the FBIJobs claim.

The recommended model is not unrestricted AI control. Leadership defines risk tolerance and preauthorizes specific, reversible, time-limited containment actions. A person remains on the loop: notified immediately, able to review and reverse the action, and accountable for escalation.

That model already exists in production form. Microsoft Defender XDR's automatic attack disruption contains a device, contains a user at the endpoint layer, or disables the account when it identifies an active attack. Every action is logged, an analyst can release the containment, and exclusions protect break-glass and service accounts. A Salesforce Transaction Security Policy blocks a large download while it is happening. An Entra risk policy forces phishing-resistant reauthentication at medium sign-in risk. Those are the preauthorized, reversible actions the products supply. The maximum duration, the immediate notification, and the named person accountable for escalation are the governance this study recommends wrapping around them, not guarantees of the products.

Preauthorized automation or AI may:

- revoke suspicious sessions and tokens;
- pause abnormal bulk downloads or API exports;
- block new MFA methods, credentials, keys, or OAuth grants;
- temporarily quarantine a device;
- temporarily suspend a high-risk identity;
- isolate a compromised workload from sensitive systems;
- preserve logs, snapshots, and volatile evidence;
- move a sensitive service into a restricted or read-only mode; and
- open an incident and notify the response team.

Each action needs high-confidence thresholds, blast-radius limits, full audit records, tested rollback, escalation timers, and regular simulation. The system should prefer reversible containment and make uncertainty visible.

Human approval should remain mandatory for:

- permanent deletion;
- broad or prolonged production shutdown;
- public communications;
- employee discipline;
- ownership or federation changes; and
- permanent security-policy changes.

Irreversible or business-state-changing actions should use dual human approval when appropriate.

Founder's Perspective: bounded speed

The business already owns the risk of responding too slowly. AI adds the risk of responding incorrectly at machine speed. Leadership has to evaluate both. When the evidence is strong and the action is reversible, pausing an export or revoking a token may be safer than waiting for a human while the attacker keeps moving. Humans should govern the boundaries and stay on the loop, not become a mandatory delay before every defensive action.

The reports from Google and Anthropic make one more point plainly. The reported operations used mainstream AI tools that employees and defenders can also use. That is not an argument against the tools. It is an argument for governing them on both sides: knowing what an attacker can do with them, and preauthorizing what a defender may do with them.

This model aligns with the NIST AI Risk Management Framework's emphasis on defined human oversight, risk-tolerance decisions, monitoring, appeal and override, incident response, and change management. It should be implemented as a governed control system, not as an assertion that AI is responsible for the outcome.

What would not have been enough

- **Annual awareness training alone:** It does not remove a vulnerable server, narrow an OAuth scope, isolate untrusted support content, or stop an export.
- **MFA alone:** It does not stop pre-authentication server exploitation. A user can also be deceived into granting an attacker-controlled app legitimate access. OTP and push methods are not phishing resistant.
- **A password reset alone:** Existing sessions, refresh tokens, OAuth grants, app passwords, or a newly enrolled factor may remain usable.
- **A WAF alone:** It does not replace endpoint removal, application isolation, host telemetry, or class-wide remediation. Google specifically cautioned against WAF-only protection for the PeopleSoft chain.
- **A compliant cloud boundary alone:** Under the shared-responsibility model, customers remain responsible for their applications, identities, data, guest systems, and configurations. A secure substrate cannot compensate for an over-privileged workload or stolen customer token.
- **Encryption at rest alone:** An attacker controlling an authorized application or identity may be able to request decryption and export through normal functions.
- **Alerts alone:** An alert that arrives after a three-minute export can support the investigation but cannot prevent the transfer.
- **Backups alone:** Backups restore availability. They do not reverse theft, doxxing, coercion, or privacy harm.
- **Indicator blocking alone:** Infrastructure, domains, tools, and filenames change. Detect the behavior chain and use indicators as supporting evidence.
- **Dual approval everywhere:** Poorly designed approval requirements can block emergency response, encourage workarounds, and create a false sense of safety.

Five executive decisions

1. **Design for the human who will eventually be deceived.** Hold people accountable for procedure, but hold the organization accountable for limiting the consequence of one mistake.
2. **Treat access as an attack surface, including public access.** Human users, service accounts, OAuth apps, support sessions, and workload identities all require least privilege, behavior baselines, and rapid revocation. Public guest permissions also need an accountable owner and recurring tests.
3. **Make high-confidence abnormal actions stop, not just alert.** Authentication, authenticator replacement, credential creation, logging changes, and bulk export deserve preplanned blocking or pause decisions.
4. **Protect the controls that protect the business.** Use two-person integrity for evidence destruction, trust-boundary changes, and permanent weakening of safeguards, with a tested emergency path.
5. **Authorize defense to operate at attacker speed.** Preapprove bounded, reversible containment and keep humans on the loop for review, escalation, and irreversible decisions.

Open questions and update triggers

This edition will be updated when any of the following occurs:

- a new FBI, DOJ, Oracle, AWS, CISA, or affected-person statement changes the FBIJobs evidence;
- any claimed FBI data is released (the one-week deadline the group set on September 22 passed without a reported release, and the group said on September 28 that it would not publish);
- the FBI, DOJ, or Oracle names the flaw behind the FBIJobs incident, publishes an affected-person count, or revises the October 5 account;
- a court record in the Dutch or Jordanian cases clarifies a participant's role in a case in this edition;
- a CVE or vendor advisory establishes or rejects the claimed new PeopleSoft path;
- a court record or criminal admission clarifies a campaign participant's role;
- an affected organization publishes a forensic report that changes initial access, scope, or attribution;
- new evidence resolves whether UNC6395 and a ShinyHunters-branded operator were connected;
- a primary source revises confirmed victim counts or affected data categories; or
- this edition is adapted for public release and requires a fresh evidence cutoff, legal review, and link validation.

Appendix A: technical and human-risk control checklist

Identity and support workflows

- ☐ Phishing-resistant MFA is required for privileged and sensitive access.
- ☐ Sensitive access requires a managed device and risk-appropriate location or network policy.
- ☐ Password reset and MFA replacement use known-good verification, with independent approval when risk is elevated.
- ☐ Password-reset response revokes sessions, refresh tokens, app passwords, OAuth grants, and suspicious authenticators as appropriate.
- ☐ Help-desk staff receive workflow-specific vishing and legitimate-screen abuse exercises.
- ☐ Suspicious calls, denied resets, and near misses are captured as telemetry, not discarded as harmless events.

SaaS, OAuth, and non-human identities

- ☐ End users cannot authorize unapproved high-impact connected apps.
- ☐ OAuth scopes, service accounts, integration tokens, and workload identities are inventoried and assigned owners.
- ☐ Tokens are short-lived where possible, rotated, isolated per customer or workload, and restricted by expected client or network.
- ☐ A tested kill switch can revoke a vendor integration across the enterprise.
- ☐ New grants, schema discovery, broad queries, and bulk exports are correlated into one risk decision.
- ☐ Support tickets and collaboration systems detect or redact secrets.

Applications, cloud, and data

- ☐ Internet-facing applications and administrative endpoints are inventoried, owned, minimized, patched, and segmented.
- ☐ Public application tiers cannot directly inherit broad HR, storage, or cloud-management access.
- ☐ Workload identities use least privilege and protected, short-lived secrets.
- ☐ Egress is deny-by-default where feasible, with detection for new destinations, tunneling, remote administration, staging, and large transfers.
- ☐ Sensitive data is classified, segmented, minimized, and retained only for a documented need.
- ☐ Abnormal downloads and API exports can be paused before completion.

Logging, governance, and response

- ☐ Identity, endpoint, help-desk, application, SaaS, database, cloud, DNS, network, and export telemetry is centralized and time synchronized.

- ☐ Logs are stored in a tamper-resistant, independently administered plane.
- ☐ Attempts to disable logging or core controls fail closed and open an incident.
- ☐ Two-person integrity covers evidence deletion, retention reduction, trust changes, mass deletion, and emergency-account modification.
- ☐ Break-glass procedures are monitored, time-limited, tested, and workable with actual staffing.
- ☐ Reversible automated containment has written thresholds, blast-radius limits, human notification, rollback, and audit records.
- ☐ Exercises test whether defenders notice employee, identity, token, and export behavior, not only whether an employee clicks.

Appendix B: prioritized 30/60/90-day action plan

First 30 days: stop the easiest repeat paths

1. Inventory privileged users, MFA methods, high-impact OAuth apps, service accounts, vendor integrations, export-capable roles, internet-facing applications, and break-glass accounts.
2. Require phishing-resistant MFA and managed devices for administrators, help-desk personnel, data-platform users, and other high-risk roles.
3. Default-deny unapproved connected-app consent and review existing grants for scope, owner, age, and business need.
4. Implement immediate revocation procedures for sessions, refresh tokens, OAuth grants, API keys, and suspicious MFA methods.
5. Identify bulk-export paths and place temporary thresholds or manual holds on the highest-risk datasets.
6. Route logging-control changes and deletion attempts to an independent, tamper-resistant alert path.
7. Preserve near-miss data from phishing, vishing, reset, enrollment, and OAuth events. Establish a non-punitive reporting channel.
8. Validate that emergency accounts work, are independently monitored, and cannot silently modify one another.

By 60 days: connect signals and approvals

1. Correlate password reset, new-factor enrollment, device posture, location, OAuth grant, token use, table discovery, and export behavior.
2. Define medium-risk and high-risk responses for each abnormal action, including explicit fail-closed thresholds.
3. Deploy or tune risk-adaptive access policies and review OAuth grants. In Entra environments, use Identity Protection and Conditional Access where supported. For Salesforce visibility, evaluate the Defender for Cloud Apps connector and the relevant Salesforce Shield Event Monitoring requirements. Test actual detection and response behavior against the licensed configuration.

4. Implement independent approval for high-risk recovery, privileged grants, new durable credentials, and abnormal export release.
5. Build a real two-person workflow for business-state-changing actions rather than assuming a single PIM approver provides dual control.
6. Segment public application tiers from sensitive data and cloud-management planes. Review the effective permissions of each workload identity.
7. Establish the vendor-integration kill switch and run a controlled exercise.
8. Deliver short, role-specific exercises for help desk, administrators, developers, support agents, executives, and high-value data users.

By 90 days: operate and test bounded defense

1. Preauthorize a first set of reversible automated actions: session and token revocation, export pause, new-factor block, device quarantine, workload isolation, evidence preservation, and restricted service mode.
2. Define confidence thresholds, maximum duration, blast radius, rollback, notification, and escalation for each action.
3. Run a purple-team exercise that begins with a vishing or stolen-token event and tests whether identity, application, and data signals combine before exfiltration completes.
4. Run a separate public-application compromise scenario that tests segmentation, workload credential exposure, egress, evidence preservation, and continuity.
5. Test the dual-control and break-glass process during limited staffing. Fix any step that depends on an unavailable person or undocumented knowledge.
6. Review retained sensitive data and remove, vault, mask, or segment records without a current business, legal, or regulatory need.
7. Present leadership with outcome metrics: attacks interrupted, exports paused, time to revoke access, near-miss trends, false-containment rate, and time to human review. Do not use training-completion percentage as the primary measure.
8. Reassess risk tolerance after the exercises and adjust automation and approval boundaries based on evidence.

Source library

2024 Snowflake customer-account campaign

- Mandiant: UNC5537 targets Snowflake customer instances
- DOJ: Connor Riley Moucka guilty plea
- DOJ: United States v. Moucka and Binns case record
- AT&T SEC filing
- TechCrunch: AT&T confirmation of the Snowflake workspace
- Live Nation SEC filing

- Ticketmaster incident notice
- WIRED: Ticketmaster, Santander, and the Snowflake campaign
- Santander incident statement
- Advance Auto Parts SEC filing
- Advance Auto Parts breach notice
- Neiman Marcus breach notice
- UK NCSC: MFA lessons from Snowflake customer breaches
- Snowflake: QUERY_HISTORY view, PUT and GET command status

2025 Salesforce and integration-token campaigns

- Google Threat Intelligence: voice phishing to data extortion
- FBI: UNC6040 and UNC6395 FLASH
- Salesforce: protecting environments from social engineering
- OAIC: report into preliminary inquiries of Qantas
- Qantas initial incident disclosure
- Google Threat Intelligence: Salesloft Drift token campaign
- Salesloft Trust Center incident chronology
- Cloudflare: response to the Salesloft Drift incident
- BleepingComputer: ShinyHunters claim and Google's response
- BleepingComputer: later reporting on the Drift operation
- CrowdStrike: PowerSchool investigation report

2026 SaaS and application campaigns

- Google Threat Intelligence: expansion of ShinyHunters SaaS data theft
- Google Threat Intelligence: defensive guidance
- Google Threat Intelligence: UNC6671 targets enterprise cloud environments
- Google Threat Intelligence: adversarial AI from prompting to autonomy
- Okta: phishing kits synchronized with caller scripts
- Instructure incident hub
- Instructure customer FAQ
- U.S. Department of Education: Canvas security alert
- FBI IC3: May 15 ShinyHunters public service announcement
- Oracle: CVE-2026-35273 advisory
- Google Threat Intelligence: PeopleSoft zero-day campaign
- CISA: Known Exploited Vulnerabilities data
- NAIC security update

Developing FBIJobs event

- 404 Media: ShinyHunters FBI claim
- Nextgov/FCW: FBIJobs data-theft claim
- POLITICO: FBI investigation statement
- BleepingComputer: claimed PeopleSoft zero-day
- CyberInsider: claimed entry point and screenshot
- TechCrunch: FBI agents and applicants data claim
- FBI: Privacy Impact Assessment for FBIJobs and Candidate Gateway
- FBIJobs: Applicant Portal FAQ

Added October 6, 2026:

- Google Threat Intelligence: renewed PeopleSoft exploitation and WAF bypass, September 25
- BleepingComputer: the WAF bypass and the actor's FBIJobs claim, September 26
- CBS News: Dutch police statement and FBI statements, September 29
- Nextgov/FCW: FBI video message after the Dutch arrest, September 29
- CyberInsider: the group's statement that it will not publish, September 30
- Reuters via US News: Jordan detention, October 3
- The National: Jordanian state media confirms an arrest, October 4
- The Record: arrests and the internal memo, October 5
- The Register: FBI confirms multiple arrests, October 5
- SecurityWeek: the FBI's statement on the unapplied patch, October 6
- The Hacker News: the same statement and the Reuters identifications, October 6

Additional primary sources

- Salesforce: Experience Cloud guest-access advisory
- FINRA: Experience Cloud alert
- Microsoft: ShinyHunters OAuth abuse
- Anthropic: September 2026 threat intelligence report

Microsoft Learn pages cited in the reference edition

- Conditional Access authentication strengths
- Token protection in Conditional Access
- Protecting tokens in Microsoft Entra, including continuous access evaluation
- Require compliant device: Conditional Access policy template
- Enable compliant network check with Conditional Access (Global Secure Access)
- Microsoft Graph: revokeSignInSessions
- Respond to a compromised cloud email account (app passwords)

- Configure how users consent to applications
- App governance: OAuth app policies
- Automatic attack disruption in Microsoft Defender

Governance and control design

- NIST SP 800-207: Zero Trust Architecture
- NIST SP 800-61 Rev. 3: Incident Response Recommendations
- NIST AI Risk Management Framework
- CISA Zero Trust Maturity Model Version 2
- Microsoft: risk-based Conditional Access
- Microsoft: Privileged Identity Management approval workflow
- Microsoft: Privileged Identity Management role settings
- AWS Shared Responsibility Model

Document control

Edition: Full report, September 23, 2026. **Last updated:** October 6, 2026.

Basis: Public statements by affected organizations, government authorities, and responders with direct incident visibility, cited inline. This is a public analysis, not a forensic report, and it confers no certification. Recommendations are defensive design judgments, not findings that a named organization lacked a listed control.

Prepared by: Mount Xion Technologies, San Antonio, Texas. mountxiontechnologies.com