theom

WHITE PAPER

Attacks at machine speed and scale

Why every door leads to the data, and how to catch AI-driven attackers there.
Lessons from the ARTEX incidents in Snowflake and Microsoft 365.

BASED ON	DOCUMENT	DATE
ARTEX incidents observed by Theom	White paper	October 2026

Executive summary

In October 2026, Microsoft Threat Intelligence reported that a state-aligned actor it tracks as Storm-0940 used ARTEX, an open-source multi-agent penetration testing framework, to compromise a software supply chain vendor and sign in to customer environments through that vendor's trusted applications. In several cases Microsoft measured less than three hours from initial targeting to privileged access.[1] The same month, investigators in South Korea found traces of ARTEX linked to breaches at several banks, where attackers came in through side-door systems and "extracted information by querying it."[2] Attackers now search for a way in at machine speed, and they arrive with valid credentials.

Theom has detected this pattern in customer environments, in Snowflake and in Microsoft 365. To preserve customer confidentiality, this paper walks through both recreated in Theom's Threat Detection Research Lab. In each, a software supply chain vendor's identity was compromised on the vendor's side, granted itself more privileges and went for the data: in Snowflake, staging and sharing member records in less than 15 minutes; in Microsoft 365, reading the mail and files of employees. Every step was a permitted operation from a valid credential, and the existing security tools did not trip.

Where Theom was set to act, the attack stopped. In one customer's production environment, Theom quarantined the vendor identity and disabled its sessions at its first unintended reach for sensitive data.

Why a data-centric behavioral defense. A typical enterprise runs hundreds of trusted vendor integrations, and any one of them can become the way in. Guarding every door is a losing game, but every door leads to the data. Attacks like these are built to look like normal use, split into pieces too small for volume alerts, so the signal is not how much an identity does but whether its behavior fits its purpose.

© Theom, 2026.

Why Theom. Theom runs where the data lives. It understands what the data is, learns what each identity and AI agent is meant to do, and acts on departures automatically, through each platform's native controls and inline for agents, within minutes of the activity being recorded. Identity, endpoint and log-monitoring tools answer whether an identity can do something. Theom answers whether it should.

> **The bottom line:** an attacker with a trusted identity now moves at machine speed and looks like normal use, but still has to reach the data. That is where it can be stopped. In both incidents, Theom saw every stage where the data lives, tied to the vendor identity behind it, and in Snowflake it had flagged the risky posture before the attack began.

01 Why assume breach now

Assume breach is an old principle. What has changed is how little time it leaves, and how creatively an attacker can now search for a way through. Two things in the Storm-0940 campaign make it the right starting point for data security planning.

- **Attackers arrive through access you already trust.** Microsoft describes the agents authenticating to customer environments through the existing trust of the vendor's cloud applications. A stolen key or token for a trusted integration is a valid login. No control at the front door is designed to reject it.
- **Agents compress the timeline and chain weaknesses creatively.** Microsoft notes that agents work around the clock, through weekends and holidays, and that the speed leaves defenders very little time to respond. Agents also try many combinations of weaknesses, often in parallel, and chain them into paths no one anticipated. In October 2026, CISA urged immediate patching of critical PHP vulnerabilities that had been found with AI.[4] A review cycle measured in days is no longer a response plan.

Prevention still matters. Key hygiene, network policies, least privilege and vendor due diligence all reduce the odds. But each of those controls can fail, and a compromised vendor is outside your control entirely. The question that remains is what happens after the attacker is in.

The data is the target

The South Korean bank breaches make the point.[2,3] The attackers came in through side systems, not the core banking platform, and took names, incomes, loan limits and resident registration numbers by querying them. Whatever the path in, an attacker after data has to query, copy, stage or share it where it is stored.

> **The planning principle:** the entry point varies from attack to attack. The datastore does not. Monitoring there, with the context to judge what the activity means, covers every path in.

Behavior, not volume

Volume alone is a weak signal. Alerts that fire when activity rises a few standard deviations above normal assume the attacker takes a lot at once. An AI-driven attack can split its work into small pieces that each stay under the threshold and look like ordinary use. What does not look ordinary is the behavior: an identity reaching

© Theom, 2026.

data it has never touched, gaining privileges it never held, connecting in a new way or moving data somewhere new.

Theom does not depend on indicator-of-compromise rules, which go stale as attackers change infrastructure and tools from one campaign to the next. Its detections measure behavior against the data, the identity's purpose and its own history. In both customer incidents, the existing security tools did not trip, because each step was a permitted operation from a valid credential.

That is the case for a data-centric behavioral defense model: judge what each identity does with the data against what it is meant to do. The model is set out in full after the two incidents that follow.

02 Two incidents, one pattern

Both incidents follow the shape Microsoft described: a trusted vendor identity in the wrong hands, escalation, then a move on the data. One played out in Snowflake and the other in Microsoft 365. Theom has seen the same pattern on other data platforms.

Screenshots recreate evidence collected in the wild; identifiers are partially redacted.

STAGE	SNOWFLAKE	MICROSOFT 365
Trusted identity	A software supply chain vendor's service account, signing in with an RSA key pair	A software supply chain vendor's Entra ID application, using app-only tokens
Normal job	Row counts and a freshness check on two tables	Scheduled backup reads of mail and files
Discover	Tables listed by row count, columns listed, an AI prompt asking which columns might hold sensitive records	1,842 Microsoft Graph requests across six directory collections, at 521 a minute
Escalate	Granted itself `ACCOUNTADMIN`	Assigned itself Global Administrator and Key Vault Secrets User
Persist	None observed	Added client secrets to two other applications
Collect	Read the member and claims tables in full	Read the mail and files of three people, more than 7,000 audit records
Move out	Unloaded 5,000 member rows to a stage and granted a share read access	Read directly through Microsoft Graph

Hygiene comes first

Assuming breach does not mean giving up on prevention. Before anyone gets in, Theom's hygiene detections clean up the conditions an attacker would rely on: service accounts with no network policy, personal access tokens with privileged roles, files reachable through anonymous links or shared with external guests, sensitive

© Theom, 2026.

data that carries no label or a weaker one than its content warrants, overprovisioned data, production data flowing into non-production, and AI agents using a person's credentials. Each one closed is one less path in, and one less place sensitive data sits unprotected when someone does get in.

Theom also runs attack simulation. It uses AI to find the paths an attacker could take to the highest-value, most sensitive data, then prioritizes those attack chains and the kill points where each one can be broken. Fixing the right weakness first removes the most exposure for the least work.

03 Snowflake: less than 15 minutes

The vendor's Snowflake service account normally ran row counts and a freshness check on two tables. At 9:43 p.m. Pacific on October 22, 2026, in less than 15 minutes, it escalated its privileges to `ACCOUNTADMIN`, listed tables by size and their columns, read the member and claims tables in full, called `claude-sonnet-4-6` through Cortex with a prompt asking it to find columns that might hold sensitive records, unloaded all 5,000 member rows to a stage and granted a share read access to the member table. Theom raised 13 findings for the identity across excessive access, weak authentication, AI activity and data movement.

One of those findings shows how a stolen credential gives itself away. A backup service account runs a fixed, programmatic job. Theom's Impersonation detection determined that this account's behavior had become AI-driven: two impersonations across 31 logins, on an identity that had reached a datastore.

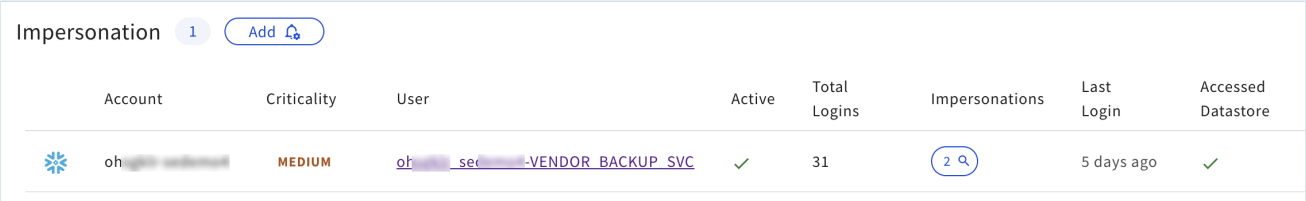


Figure 1. The Impersonation finding for the vendor's service account: an identity built for a fixed programmatic job, now driven by AI.

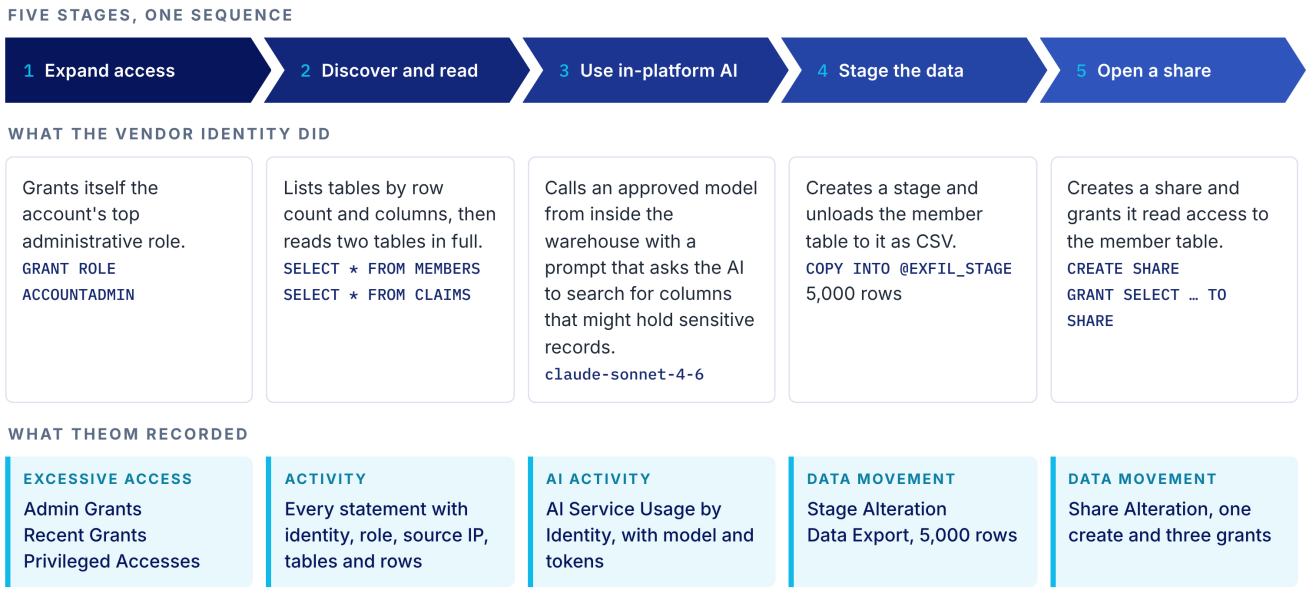


Figure 2. The five stages of the sequence and the Theom findings each one produced.

© Theom, 2026.

The weakness was visible before the attack

When the vendor was onboarded, the service account was authorized for `MANAGE GRANTS` for legitimate purposes. Theom flagged it as a medium-criticality finding, and flagged the same identity as high criticality for having no network policy. Neither meant the vendor was misbehaving. Together they described a trusted identity that could expand its own access from anywhere, one grant away from the most sensitive data: the path Theom's attack simulation flagged, and the cue to add compensating controls before anything went wrong.

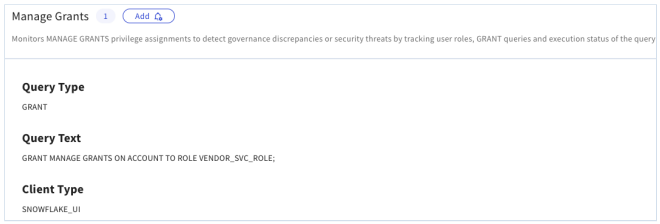


Figure 3. Theom's Manage Grants finding for the vendor role, recorded before the attack began. A necessary privilege, but one that calls for compensating controls.

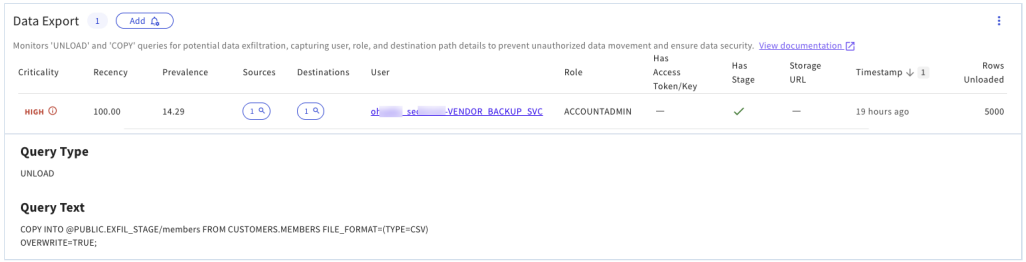


Figure 4. The Data Export finding: the vendor identity, acting as `ACCOUNTADMIN`, unloaded all 5,000 rows of the member table to a stage, with the exact query attached.

04 Microsoft 365: from directory to mailbox

The software supply chain vendor's Entra ID application normally read mail and files on a backup schedule. Once in the wrong hands, it mapped the tenant through Microsoft Graph, assigned itself Global Administrator and Key Vault Secrets User, added client secrets to two other applications and read the mail and files of three people with an app-only token. Theom raised nine findings across four detections, each keyed to the application's service principal.

The trust was the exposure

A backup application is granted tenant-wide rights to read mail and files because that is its job. For the self-grants in this incident to succeed, the application also had to hold rights to assign roles. Together those describe a trusted identity that could expand its own access and read anyone's mail, which is exactly the kind of identity an administrator should scrutinize and wrap in compensating controls, such as conditional access that restricts where the application can sign in from, before anything goes wrong. It is the path Theom's AI flags in its attack simulation: a trusted application one role assignment away from everyone's mail.

© Theom, 2026.

FOUR STAGES, ONE TRUSTED APPLICATION

1 Discover the directory | 2 Grant itself more | 3 Plant backdoors | 4 Read mail and files

WHAT THE VENDOR APPLICATION DID

1 Discover the directory	2 Grant itself more	3 Plant backdoors	4 Read mail and files

Enumerates users, groups, applications, service principals, directory roles and sites through Microsoft Graph.
`1,842 requests`
`521 per minute`
	Assigns itself Global Administrator in Entra ID and Key Vault Secrets User on a resource group.
	`Add member to role`
	`roleAssignments/write`
		Adds new client secrets to two other applications, so access survives a rotation of its own credential.
		`Secrets valid to 2028`
			Uses an app-only token, with no user behind it, to read the mail and files of three people.
			`7,179 audit records`
			`MailItemsAccessed`
			`FileDownloaded`

WHAT THEOM RECORDED

RECONNAISSANCE	INFILTRATION	INFILTRATION	DATA GATHERING
Machine-Speed Directory Recon, two findings	Service Principal Self-Grant, two findings	App Credential Added by App, two findings	App-Only Mailbox & File Access, three findings

Figure 5. The four stages of the chain and the Theom detection that recorded each one.

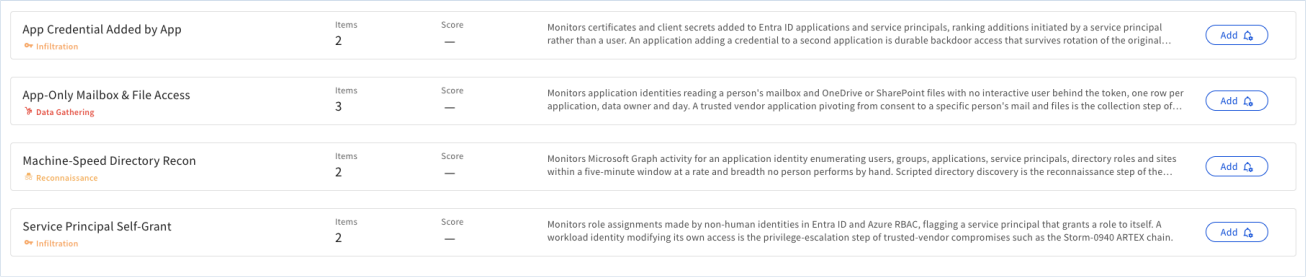


Figure 6. The four detections that recorded the chain, each classified by the stage it represents.

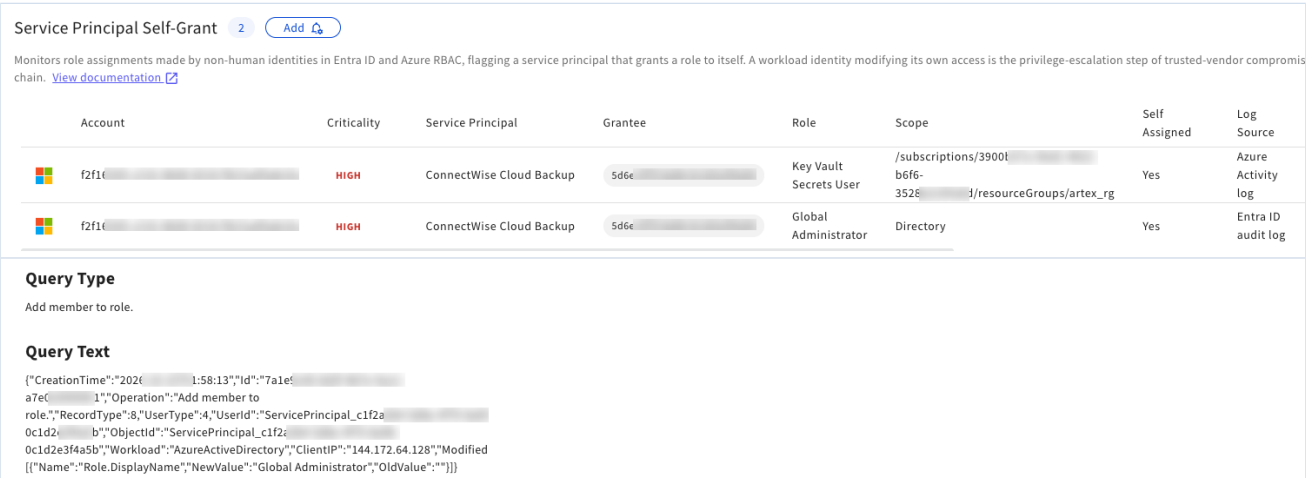


Figure 7. The self-grants: the vendor application made itself Global Administrator and gave itself a Key Vault secrets role. The audit record underneath shows the role change.

© Theom, 2026.

App-Only Mailbox & File Access 3 Add

Monitors application identities reading a person's mailbox and OneDrive or SharePoint files with no interactive user behind the token, one row per application, data owner and day. A trusted vendor application pivoting from consent to a specific person's mail and files is the collection step of the Storm-0940 ARTEX chain. View documentation

Data Owner	Access Kind	Audit Records	Access Summary	Workloads	Source IPs	User Agents	First Seen	Last Seen	App Id	Is Malicious	Operations	Detai
test-user-1@	Files	3377	0 mailbox access records, 3377 file access records	onedrive, sharepoint	144.172.64.128	azure-toolkit/1.0	10 days ago	10 days ago	50 - 6e9 - 942b- a099	⚠	FileDownloaded, FileSyncDownloadedFull	
test-user-2@	Mail	2106	2106 mailbox access records, 0 file access records	exchange	144.172.64.128	azure-toolkit/1.0	10 days ago	10 days ago	50 - 6e9 - 942b- a099	⚠	MailItemsAccessed	
test-user-3@	Mail + Files	1696	1284 mailbox access records, 412 file access records	exchange, onedrive	144.172.64.128	azure-toolkit/1.0	10 days ago	10 days ago	50 - 6e9 - 942b- a099	⚠	FileDownloaded, MailItemsAccessed	

Figure 8. The collection step: one finding per data owner, 7,179 audit records in all, from an address flagged as malicious, with the Exchange and OneDrive operations behind each.

05 Every step was permitted

Nothing in either incident required an exploit. The credentials were valid. The self-grants worked because each vendor identity already held rights to grant roles. Full-table reads, unloads and shares in Snowflake, and directory, mail and file reads through Microsoft Graph, are operations those identities were permitted to perform. Access control answered the question it is built for, whether this identity can, and the answer was yes each time.

The question that matters after a breach is different: should this identity be using this data in this way, for this purpose? Three kinds of context answer it.

- **The data.** What each table, mailbox and file holds, how sensitive and valuable it is, and where it flows. A member table holds names, dates of birth, member IDs and plan details, and a mailbox holds contracts and customer correspondence, much like the records taken from the Korean banks. Theom classifies them, so every finding is prioritized by the sensitivity of the data it touched.
- **The identity and its purpose.** A backup integration has a known job. Self-elevation, full-table reads, unloads, new shares, planted secrets and on-demand reads of specific people's mail fall outside it whatever privileges the identity holds.
- **The behavior.** Years of near-identical backup activity, then a burst of grants, reads and movement. The change is the signal.

The evidence for both chains existed, but spread across separate logs and tools. A security team could have rebuilt the sequence after the fact. By then the data would already have left.

> **The distinction that decides it:** a permitted operation can still be unintended use. Telling the two apart takes knowledge of the data, the identity's purpose and its observed behavior, held together.

© Theom, 2026.

06 A data-centric behavioral defense model

These incidents suggest an operating model for the period after an identity is compromised. It judges behavior against the data rather than counting activity. It has four steps, and each one depends on the step before it.

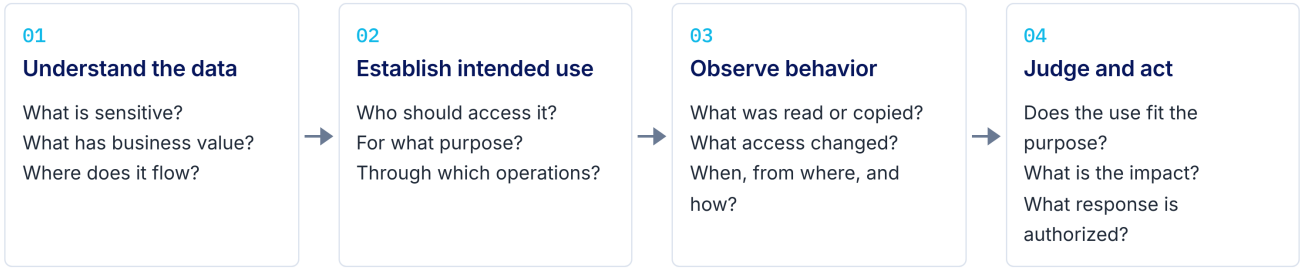


Figure 9. The data-centric behavioral defense model.

In Theom's terms, understanding the data means classifying it in place and tracing where it flows. Observing use means reading each platform's own activity record continuously, query logs in Snowflake and audit, sign-in and Graph activity logs in Microsoft 365, with every read, grant, credential change and data movement tied to the identity behind it. Judging it means ranking departures by the value of the data involved and handing responders the evidence. The step that most programs skip is the second one.

Each attacker objective leaves a trace where the data lives, and hygiene removes many of the openings before an attacker arrives. The table maps a selection of Theom's detections on each platform to each stage.

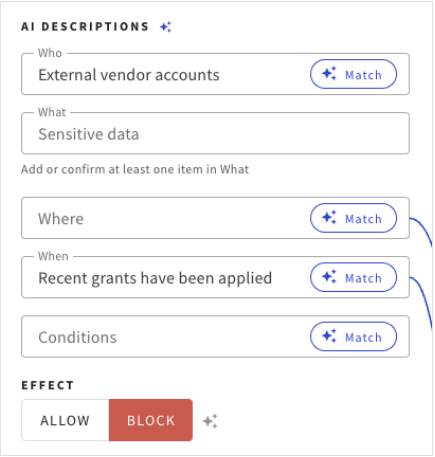
ATTACKER OBJECTIVE	SNOWFLAKE	MICROSOFT 365
Before the attack	Service Accounts without Network Policy · PAT Inventory · Unlabeled Sensitive Data · Masking Policy Conflicts · Public Accessible Data Objects	Anonymous Link Exposure · Org-Wide Data Exposure · External Guest Access · Under-Labeled Sensitive Data · Overprovisioned Data Assets
Get in and discover	Failed Logins · Logins from Malicious IP Addresses · Atypical Login Events · Impersonation	Machine-Speed Directory Recon
Dig in	Admin Grants · Recent Grants · Manage Grants · Atypical Grants to Objects · Network Policy Modification	Service Principal Self-Grant · App Credential Added by App
Collect	Privileged Accesses · AI Service Usage by Identity · Atypical Read Events · Secrets Exposure	App-Only Mailbox & File Access
Get it out or cover tracks	Data Export · Stage Alteration · Share Alteration · Atypical Download/Copy	Sensitivity Label Downgrades · Destructive Data Activity

© Theom, 2026.

Establishing intended use: codified and learned

Intended use is rarely written down, which is why this step stalls. Theom establishes it in two ways, and each covers the other's blind spot.

Codified. The business states what good looks like in plain language, and Theom turns it into a policy. A security team worried about this exact scenario might write: external vendor accounts should not be able to access sensitive data if they have had a recent change in their access rights. Theom resolves that sentence into who, what, where, when and an effect, and the same policy model applies across data platforms and agent frameworks. At the data layer Theom enforces through each platform's native controls, such as Purview in Microsoft 365; at the agent layer it acts on the request inline, including for agents built in Copilot Studio. This is the policy that stops both chains: it describes the moment each vendor identity's access changed and it went for the data.





From the sentence, Theom proposes each part of the policy:

- **Who:** external vendor accounts
- **What:** sensitive data, resolved through classification
- **Where:** where the data is stored, for residency and geofencing; where the access comes from, such as a user who works from New York suddenly connecting from Germany; and where the data came from, such as a production copy that has landed in development
- **When:** recent grants have been applied, and the time of day or week, such as an agent that does not run on weekends
- **Effect:** block

An administrator matches each part to real principals, data types and accounts, then saves the policy.

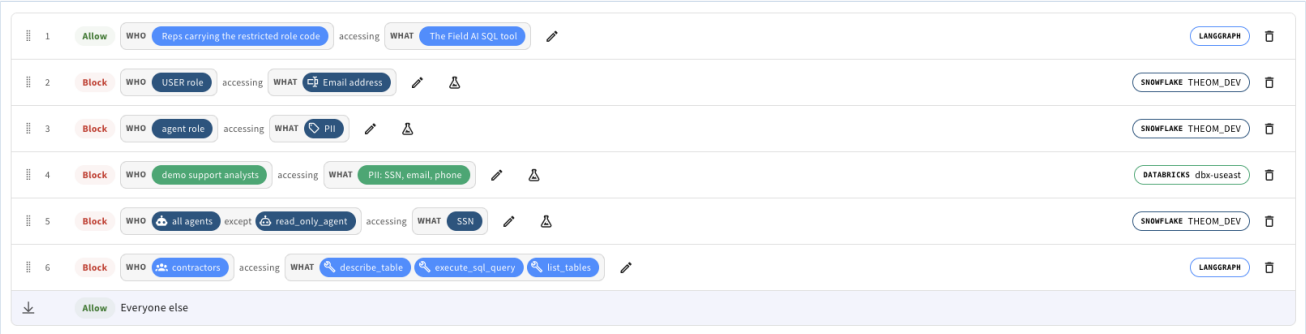


Figure 10. Writing policy in plain language. A sentence becomes who, what, where, when and an effect, and one ordered policy set covers Snowflake, Databricks and agent tools in a LangGraph chain.

Learned. Theom also builds baselines from observed activity for each identity and flags departures from them, even where no one has written a rule. Detections keyed on rare events, such as self-grants, staging, sharing, planted credentials and machine-speed reconnaissance, fire from the first occurrence, with no history needed. For backup integrations with years of history behind them, full-table reads, an unload or on-demand reads of specific people's mail are exactly the departures a baseline exists to surface.

© Theom, 2026.

Guardrails for AI identities. Theom learns from intended behavior and automatically establishes data guardrails for AI identities. Even when an AI agent acts under a human's identity, Theom knows it is AI and holds it to the guardrails meant for it, so an agent that strays from its purpose is stopped at the data.

The two halves need each other. Codified policy covers new integrations and agents from day one, before there is any history to learn from. Learned baselines catch what nobody thought to write down.

07 Responding at the speed of the attack

Detection matters only if it leads to an authorized action quickly. Theom reads each platform's own activity record continuously and alerts in minutes, not at the next review. It stays out of the query path and acts through each platform's native controls, such as Purview in Microsoft 365, with an impact analysis in Snowflake before any change is applied. At the AI layer, where enabled, Theom can allow, block, mask or redact an AI request as it is made, including from Copilot Studio agents.

Where the chain stops, and how fast

Customers choose, by environment and by stage, where Theom acts and where it observes. In one customer's production environment, the policy quarantined the vendor identity and disabled its sessions as soon as it departed from intended use. Environments set to observe only get a complete record of what the identity touched, but no intervention. The chain can be cut at the privilege escalation, the first access to sensitive data or the creation of an export stage.

In automatic mode no person has to approve the response: Theom can open a ticket, raise an alert, act through native controls or apply an agent policy inline. Reaction time is bounded by how quickly each platform records activity. Theom acts on Snowflake events within one to two minutes, and Microsoft Graph activity logs arrive with five to ten minutes of latency. Against attacks that run in less than 15 minutes, both leave time to act. A review queue does not.

What the response looks like

Where the response was set to act, the guardrails stopped each vendor identity at its first unintended reach for sensitive data. The steps behind that response:

For the Snowflake service account. Revoke `ACCOUNTADMIN`, disable the user or rotate its key, and pin it to the vendor's addresses. Drop the share and the stage. Establish which tables, columns and rows were read, unloaded or shared. Keep the grant rights only with compensating controls, and separate administration from data access.

For the Microsoft 365 application. Disable the service principal and revoke its credentials and tokens. Delete the self-assigned role assignments and rotate secrets the Key Vault role could reach. Remove the secrets it planted on the two other applications. Use the App-Only Mailbox & File Access findings to establish whose mail and files were read, then limit the application's mail access to the mailboxes it backs up.

Responders can work at machine speed too. An AI agent working through Theom's MCP interface reconstructed each chain in this paper from Theom's findings, without querying Snowflake or Microsoft 365 directly.

© Theom, 2026.

Six things to do now

- **Inventory and classify your data.** Know what you hold, where it lives and how sensitive it is, so you protect the right data and act on the right findings first.
- **Inventory non-human identities.** List every service account, vendor integration and AI agent that can reach your data platforms and Microsoft 365, and what data each can reach.
- **Baseline their purpose.** Record the objects, operations and volumes each one needs, and turn it into policy. Let baselines learn the rest.
- **Scrutinize grant rights.** Some integrations need elevated privileges to do their job. Treat each as a high-risk identity and pin it to the vendor's network addresses, so a stolen key only works from there.
- **Plant tripwires.** Place decoy datasets that no legitimate identity has a reason to touch. Any access to one is a high-confidence signal.
- **Agree the response in advance.** Route the high-signal findings, such as self-grants, planted credentials, unloads, new shares and app-only reads of mail, to the team that can act, with containment already approved and automated where you choose.

Start with the one platform where the most sensitive data lives, then expand.

> **The bottom line:** an attacker with a trusted identity can now move at machine speed. Where the data lives is where that speed runs out of cover, because every read, grant, credential change and data movement is recorded there. Understand the data, establish what each identity is for, and the departure from intended use becomes visible in time to act.

Your Data. Your AI. Used as Intended.

© Theom, 2026.

About Theom

Theom is data and AI security and governance that lives inside your data. Running natively inside platforms like Snowflake and Databricks, Theom sees who and what touches sensitive data, from people to applications to AI agents, and enforces policy so enterprises can put their data to work in AI safely. Customer data stays in the customer's own environment. Learn more at theom.ai.

Sources

1. Microsoft Threat Intelligence, "Storm-0940 uses ARTEX agentic AI in ConnectWise supply chain compromise," Microsoft Security Blog, October 21, 2026. microsoft.com/en-us/security/blog/2026/10/21/storm-0940-uses-artex-agentic-ai-connectwise-supply-chain-compromise
2. The Korea Herald, "Customer information leaked through loan inquiry and employee support systems at major banks," October 2026. koreaherald.com/article/10892290
3. The Herald Business, "Exclusive: Chinese AI tool ARTEX used in wave of bank hacks, probe finds," October 2026. mbiz.heraldcorp.com/article/10892497
4. CISA, "Critical Vulnerabilities in PHP Require Immediate Patching," October 21, 2026. cisa.gov/news-events/alerts/2026/10/21/critical-vulnerabilities-php-require-immediate-patching

© Theom, 2026.