The Rise of MaaS & Lumma Info Stealer

By Emily Megan Lim

Published: 2023-09-06 · Archived: 2026-04-10 02:05:42 UTC

What “securing AI” actually means (and doesn’t)

Security teams are under growing pressure to “secure AI” at the same pace which businesses are adopting it. But in many organizations, adoption is outpacing the ability to govern, monitor, and control it. When that gap widens, decision-making shifts from deliberate design to immediate coverage. The priority becomes getting something in place, whether that’s a point solution, a governance layer, or an extension of an existing platform, rather than ensuring those choices work together.

At the same time, AI governance is lagging adoption. 37% of organizations still lack AI adoption policies, shadow AI usage across SaaS has surged, and there are notable spikes in anomalous data uploads to generative AI services.

First and foremost, it’s important to recognize the dual nature of AI risk. Much of the industry has focused on how attackers will use AI to move faster, scale campaigns, and evade detection. But what’s becoming just as significant is the risk introduced by AI inside the organization itself. Enterprises are rapidly embedding AI into workflows, SaaS platforms, and decision-making processes, creating new pathways for data exposure, privilege misuse, and unintended access across an already interconnected environment.

Because the introduction of complex AI systems into modern, hybrid environments is reshaping attacker behavior and exposing gaps between security functions, the challenge is no longer just having the right capabilities in place but effectively coordinating prevention, detection, investigation, response, and remediation together. As threats accelerate and systems become more interconnected, security depends on coordinated execution, not isolated tools, which is why lifecycle-based approaches to governance, visibility, behavioral oversight, and real-time control are gaining traction.

From cloud consolidation to AI systems what we can learn

We have seen a version of AI adoption before in cloud security. In the early days, tooling fragmented into posture, workload/runtime, identity, data, and more. Gradually, cloud security collapsed into broader cloud platforms. The lesson was clear: **posture without runtime misses active threats; runtime without posture ignores root causes.** Strong programs ran both in parallel and stitched the findings together in operations.

Today’s AI wave stretches that lesson across *every* domain. Adversaries are **compressing “time-to-tooling”** using LLM-assisted development (“vibecoding”) and recycling public PoCs at unprecedented speed. That makes it difficult to secure through siloed controls, because the risk is not confined to one layer. It emerges through interactions across layers.

Keep in mind, most modern attacks don't succeed by defeating a single control. They succeed by moving through the gaps between systems faster than teams can connect what they are seeing. Recent exploitation waves like React2Shell show how quickly opportunistic actors operationalize fresh disclosures and chain misconfigurations to monetize at scale.

In the React2Shell window, defenders observed rapid, opportunistic exploitation and iterative payload diversity across a broad infrastructure footprint, strains that outpace signature-first thinking.

You can stay up to date on attacker behavior by signing up for our newsletter where Darktrace's threat research team and analyst community regularly dive deep into threat finds.

Ultimately, speed met scale in the cloud era; AI adds interconnectedness and orchestration. Simple questions — *What happened? Who did it? Why? How? Where else?* — now cut across identities, SaaS agents, model/service endpoints, data egress, and automated actions. The longer it takes to answer, the worse the blast radius becomes.

The case for a platform approach in the age of AI

Think of security fusion as the connective tissue that lets you **prevent, detect, investigate, and remediate in parallel**, not in sequence. In practice, that looks like:

1. **Unified telemetry with behavioral context** across identities, SaaS, cloud, network, endpoints, and email —so an anomalous action in one plane automatically informs expectations in others. (Inside-the-SOC investigations show this pays off when attacks hop fast between domains.)
2. **Pre-CVE and "in-the-wild" awareness** feeding controls *before* signatures—reducing dwell time in fast exploitation windows.
3. **Automated, bounded response** that can contain likely-malicious actions at machine speed *without* breaking workflows—buying analysts time to investigate with full context. (Rapid CVE coverage and exploit-wave posts illustrate how critical those first minutes are.)
4. **Investigation workflows that assume AI is in the loop**—for both defenders and attackers. As adversaries adopt "agentic" patterns, investigations need graph-aware, sequence-aware reasoning to prioritize what matters early.

This isn't theoretical. It's reflected in the Darktrace posts that consistently draw readership: **timely threat intel with proprietary visibility** and **executive frameworks** that transform field findings into operating guidance.

The five questions that matter (and the one that matters more)

When alerted to malicious or risky AI use, you'll ask:

1. **What happened?**
2. **Who did it?**

3. **Why did they do it?**
4. **How did they do it?**
5. **Where else can this happen?**

The sixth, more important question is: **How much worse does it get while you answer the first five?** The answer depends on whether your controls operate in sequence (slow) or in **fused parallel** (fast).

What to watch next: How the AI security market will likely evolve

Security markets tend to follow a familiar pattern. New technologies drive an initial wave of specialized tools (posture, governance, observability) each focused on a specific part of the problem. Over time, those capabilities consolidate as organizations realize the new challenge is coordination.

AI is accelerating the shift of focus to coordination because AI-powered attackers can move faster and operate across more systems at once. Recent exploitation waves show exactly this. Adversaries can operationalize new techniques and move across domains, turning small gaps into full attack paths.

Anticipate a continued move toward more integrated security models because fragmented approaches can't keep up with the speed and interconnected nature of modern attacks.

Building the Groundwork for Secure AI: How to Test Your Stack's True Maturity

AI doesn't create new surfaces as much as it **exposes the fragility of the seams that already exist**.

Darktrace's own public investigations consistently show that modern attacks, from LinkedIn-originated phishing that pivots into corporate SaaS to multi-stage exploitation waves like BeyondTrust CVE-2026-1731 and React2Shell, succeed not because a single control failed, but because no control saw the whole sequence, or no system was able to respond at the speed of escalation.

Before thinking about "AI security," customers should ensure they've built a security foundation where visibility, signals, and responses can pass cleanly between domains. That requires pressure-testing the seams.

Below are the key integration questions and stack-maturity tests every organization should run.

1. Do your controls see the same event the same way?

Integration questions

- When an identity behaves strangely (impossible travel, atypical OAuth grants), does that signal automatically inform your email, SaaS, cloud, and endpoint tools?
- Do your tools normalize events in a way that lets you correlate identity → app → data → network without human stitching?

Why it matters

Darktrace's public SOC investigations repeatedly show attackers **starting in an unmonitored domain, then pivoting into monitored ones**, such as phishing on LinkedIn that bypassed email controls but later appeared as anomalous SaaS behavior.

If tools can't share or interpret each other's context, AI-era attacks will outrun every control.

Tests you can run

1. Shadow Identity Test

- Create a temporary identity with no history.
- Perform a small but unusual action: unusual browser, untrusted IP, odd OAuth request.
- Expected maturity signal**:** other tools (email/SaaS/network) should immediately score the identity as high-risk.

2. Context Propagation Test

- Trigger an alert in one system (e.g., endpoint anomaly) and check if other systems automatically adjust thresholds or sensitivity.
- Low maturity signal: nothing changes unless an analyst manually intervenes.

2. Does detection trigger coordinated action, or does everything act alone?

Integration questions

- When one system blocks or contains something, do other systems automatically tighten, isolate, or rate-limit?
- Does your stack support bounded autonomy — automated micro-containment without broad business disruption?

Why it matters

In public cases like BeyondTrust CVE-2026-1731 exploitation, Darktrace observed rapid C2 beaconing, unusual downloads, and tunneling attempts across multiple systems. Containment windows were measured in minutes, not hours.

Tests you can run

1. Chain Reaction Test

- Simulate a primitive threat (e.g., access from TOR exit node).
- Your identity provider should challenge → email should tighten → SaaS tokens should re-authenticate.
- Weak seam indicator: only one tool reacts.

2. Autonomous Boundary Test

- Induce a low-grade anomaly (credential spray simulation).
- Evaluate whether automated containment rules activate without breaking legitimate workflows.

3. Can your team investigate a cross-domain incident without swivel-chairing?

Integration questions

- Can analysts pivot from identity → SaaS → cloud → endpoint **in one narrative**, not five consoles?
- Does your investigation tooling use **graphs or sequence-based reasoning**, or is it list-based?

Why it matters

Darktrace's Cyber AI Analyst and DIGEST research highlights why investigations must interpret structure and progression, not just standalone alerts. Attackers now move between systems faster than human triage cycles.

Tests you can run

1. One-Hour Timeline Build Test

- Pick any detection.
- Give an analyst one hour to produce a full sequence: entry → privilege → movement → egress.
- Weak seam indicator: they spend >50% of the hour stitching exports.

2. Multi-Hop Replay Test

- Simulate an incident that crosses domains (phish → SaaS token → data access).
- Evaluate whether the investigative platform auto-reconstructs the chain.

4. Do you detect intent or only outcomes?

Integration questions

- Can your stack detect the setup behaviors before an attack becomes irreversible?
- Are you catching pre-CVE anomalies or post-compromise symptoms?

Why it matters

Darktrace publicly documents multiple examples of pre-CVE detection, where anomalous behavior was flagged days before vulnerability disclosure. AI-assisted attackers will hide behind benign-looking flows until the very last moment.

Tests you can run

1. Intent-Before-Impact Test

- Simulate reconnaissance-like behavior (DNS anomalies, odd browsing to unknown SaaS, atypical file listing).
- Mature systems will flag *intent* even without an exploit.

2. CVE-Window Test

- During a real CVE patch cycle, measure detection lag vs. public PoC release.
- Weak seam indicator: your detection rises only after mass exploitation begins.

5. Are response and remediation two separate universes?

Integration questions

- When you contain something, does that trigger root-cause remediation workflows in identity, cloud config, or SaaS posture?
- Does fixing a misconfiguration automatically update correlated controls?

Why it matters

Darktrace's cloud investigations (e.g., cloud compromise analysis) emphasize that remediation must close both runtime and posture gaps in parallel.

Tests you can run

1. Closed-Loop Remediation Test

- Introduce a small misconfiguration (over-permissioned identity).
- Trigger an anomaly.
- Mature stacks will: detect → contain → recommend or automate posture repair.

2. Drift-Regression Test

- After remediation, intentionally re-introduce drift.
- The system should immediately recognize deviation from known-good baseline.

6. Do SaaS, cloud, email, and identity all agree on "normal"?

Integration questions

- Is "normal behavior" defined in one place or many?
- Do baselines update globally or per-tool?

Why it matters

Attackers (including AI-assisted ones) increasingly exploit misaligned baselines, behaving “normal” to one system and anomalous to another.

Tests you can run

1. Baseline Drift Test

- Change the behavior of a service account for 24 hours.
- Mature platforms will flag the deviation early and propagate updated expectations.

2. Cross-Domain Baseline Consistency Test

- Compare identity’s risk score vs. cloud vs. SaaS.
- Weak seam indicator: risk scores don’t align.

Final takeaway

Security teams should ask be focused on how their stack operates **as one system before AI amplifies pressure on every seam.**

Only once an organization can reliably detect, correlate, and respond **across domains** can it safely begin to secure AI models, agents, and workflows.
