The Magazine for Secure Applications & APIs

ANNUAL REPORT · VOL. I · EDITION 2026

PROPHAZE

EXPRESS

AI-NATIVE WAAP PLATFORM

100+
Customers

100K+
Apps & APIs Secured

1B+
Requests Daily

50M+
Threats Blocked Daily

25+
Countries

15 Min
Deployment

Up to 70%
Lower Security TCO

THREAT RESEARCH

The AI, API & Application Security Landscape Report 2026

THREAT LANDSCAPE
Q2 2026 Attack Trends Impacting Applications and APIs

AI-POWERED DEFENSE
Agentic AI, ML & Bot Protection for Zero Trust Security

CLOUD-NATIVE EDGE
Securing Modern Apps Across Cloud, Containers & Kubernetes

Prophaze™
The New Phase of Security

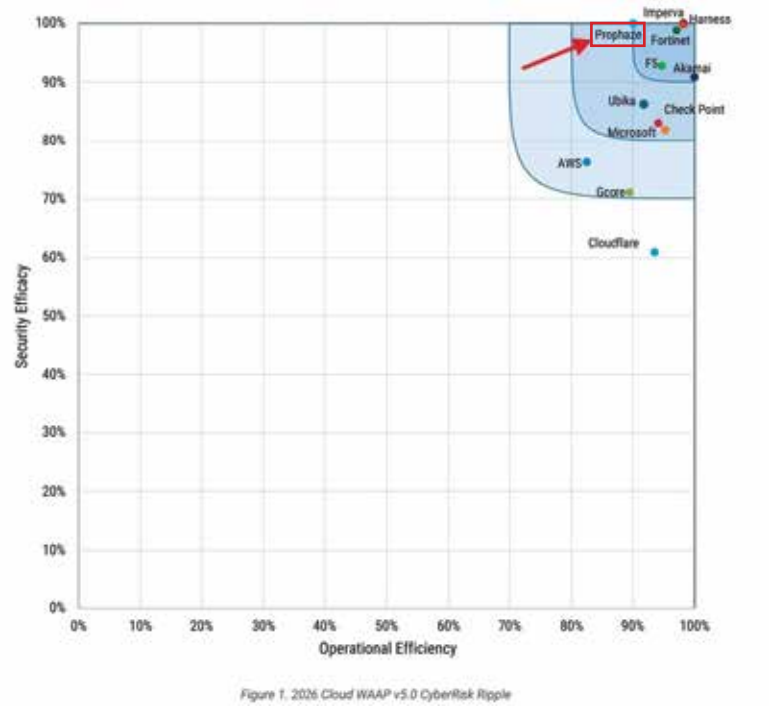


Figure 1. 2026 Cloud WAAP v5.0 CyberRisk Ripple

Exciting News!

We’re proud to announce that Prophaze has been recognized as a Leader in the SecureIQLab Cloud WAAP v5.0 CyberRisk Validation 2026!

This recognition is a strong validation of our WAAP capabilities and a milestone we can all be proud of.
A special thanks to everyone whose dedication and hard work made this achievement possible!

Prophaze™
The New Phase of Security

Unified Security for **API, App & Agentic AI**

2026 SecureIQLab WAAP Leader

Vendor in GARTNER Market Guide WAAP 2026

GARTNER Peer Insights STRONGER PERFORMER Voice of the Customer 2025

Vendor in GARTNER Market Guide WAAP 2025

Mentioned in GARTNER Market Guide API Security 2024

CONTENT

06 Executive Summary

From perimeter to interface

07 2026 by the Numbers

Six vectors, sourced and moving wrong

08 Why Security Changed

Seven epochs of expanding surface

CYBERSECURITY

10 The 2026 Threat Landscape

Prompt injection to agent abuse

AI SECURITY

14 Lessons from the Breaches

Five case files, 2022–2025

THREAT INTEL

16 Industry Risk Matrix

Sectors ranked by breach cost

RISK & COMPLIANCE

17 Maturity: Where Enterprises Stand

Reactive to autonomous defense

ENTERPRISE READINESS

18 The Modern Defense

Cloud-native, runtime, discovery

CLOUD-NATIVE

19 Frameworks & Roadmap

The six-step operating loop

GOVERNANCE

20 Outlook: Preparing for 2027

Five forecasts for 2026–2027

INDUSTRY OUTLOOK

FROM PERIMETER TO INTERFACE

BY PROPHAZE THREAT RESEARCH

The perimeter is now an interface. Every AI feature ships as an API, every API extends the attack surface, and every enterprise now defends software it never fully sees. (Prophaze Threat Research, 2026 Outlook)

Two decades of architectural change (cloud, containers, microservices, APIs, and now generative and agentic AI) have collapsed the network perimeter into a web of programmable interfaces. Each shift added an attack surface; none removed the layer beneath it. The result is a single, connected surface where an AI prompt is an API request, a model call is an endpoint, and an autonomous agent acts with credentials that are either governed or invisible.

This report synthesizes primary-source research from across the industry into an operating picture of the 2026 threat landscape and a board-measurable roadmap for reducing exposure. Six findings define the year ahead:

- **APIs are multiplying faster than they can be governed.** API estates grew 167% year over year, while only 7.5% of organizations run a dedicated API threat-modeling program.
- **Attacks are following the interfaces.** Web application and API attacks rose 33% year over year, and 37% of all web traffic is now bad-bot activity.

- **AI has become its own attack surface.** Prompt injection, tool and agent abuse, poisoning, and sensitive-data leakage are now mapped by OWASP LLM Top 10 v2.0 and MITRE ATLAS, and, as Section 03 shows, are already producing real production incidents, not just theoretical risk.

- **Shadow AI is entering through the API layer.** 78% of AI users bring their own tools to work, bypassing procurement, security, and legal review.

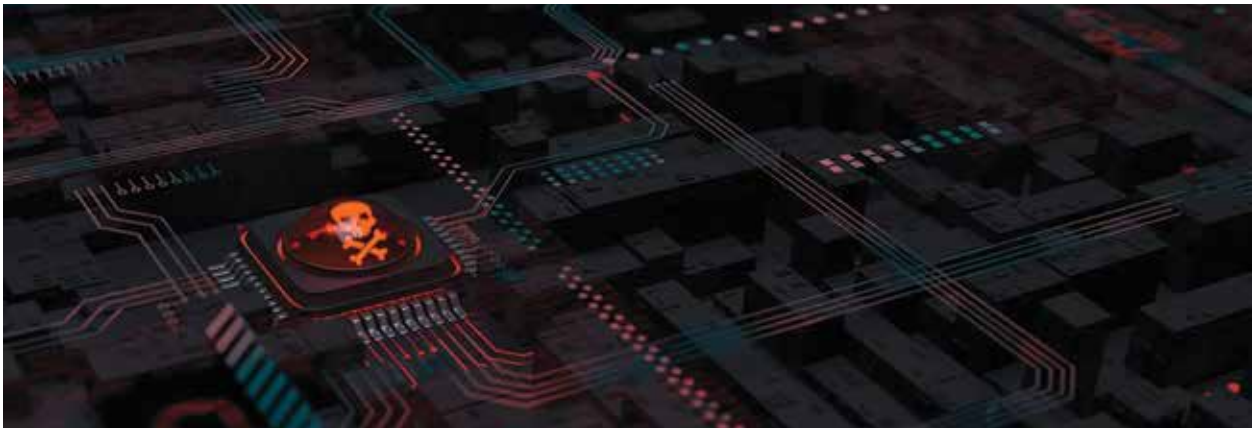
- **The failures are fundamentals, not exotics.** Unauthenticated endpoints, forged tokens, un-inventoried third-party APIs, and missing MFA drove the largest breaches of 2022–2025.

- **Speed has moved the decision point to runtime.** The median cloud attack chain now unfolds in under ten minutes, while the mean breach lifecycle remains 258 days.

2026 by the Numbers

Six measurable vectors define the modern application attack surface. Each is sourced, and each is moving in the wrong direction.

Metric	What it measures	Source
167%	Year-over-year growth in APIs under management	Salt Security, 2024
33%	Year-over-year growth in web application & API attacks	Akamai SOTI, 2025
37%	Share of all web traffic that is bad-bot activity	Imperva Bad Bot Report, 2025
78%	AI users who bring their own tools to work (Shadow AI)	Microsoft Work Trend Index, 2024
7.5%	Organizations running a dedicated API threat-modeling program	Salt Security, 2024
258 days	Mean breach lifecycle (time to identify and contain)	IBM Cost of a Data Breach



Why Security Changed: From Perimeter to Interface

The modern application is a seven-layer system in which every hop is an interface and every interface is a control point. Understanding how the surface got this way is the first step to defending it.

Seven epochs of expanding surface

Each architectural era added a new class of exposure without retiring the last. Perimeter defenses built for one era do not translate to the next.

Era	Epoch	What it introduced
2005	Traditional Apps	Monolithic web, defended by a perimeter WAF
2010	Cloud	Elastic infrastructure, shared identity
2015	Containers	Immutable, ephemeral workloads
2017	Microservices	Service mesh and east-west traffic
2020	APIs	Interfaces become the product
2023	AI	LLMs, RAG, and vector data
2025 →	Agentic AI	Autonomous actors and the Model Context Protocol (MCP)

Sources: CNCF Annual Cloud Native Survey 2025 · Gartner (agentic AI in 33% of enterprise apps by 2028) · OWASP LLM Top 10 v2.0 (2024) · MITRE ATLAS (2024 update).

The connected attack surface

Six overlapping attack surfaces (web, API, cloud, container, AI/LLM, and identity) now converge on a single enterprise. The perimeter is no longer a line; it is a graph. Every asset is either governed (documented and monitored) or ungoverned (shadow, unknown, or drifting from policy). The governing principle is unchanged and unforgiving: what you cannot see, you cannot secure.

An AI application concentrates this problem. Each of its components (the prompt interface, the model endpoint, retrieval and embeddings, training and fine-tuning data, tools and agents, and output channels) is simultaneously an authentication boundary and a data-classification decision. Left unguarded, they expose the enterprise to prompt injection, schema abuse, broken authentication, embedding leakage, retrieval poisoning, model poisoning, unbounded consumption, excessive agency, and sensitive-information disclosure.

Why AI Security Needs More Than API Security

Every AI feature is ultimately delivered through an API, and the Model Context Protocol makes that explicit: MCP is, in effect, an API standard that lets an LLM call external tools and data sources. API security is therefore the foundation AI security is built on: authentication, schema validation, and rate limiting still have to hold at every model and tool endpoint, or nothing built above them matters.

But a request can be syntactically perfect and still be malicious. An API gateway can confirm a call is well-formed, authenticated, and within rate limits; it cannot tell whether the natural-language prompt riding inside that call is attempting to override system instructions, exfiltrate data through a tool call, or steer an agent toward an unauthorized action. Closing that gap is what separates AI security from API security applied to a new endpoint type: it requires a semantic and contextual analysis layer, one that inspects prompt intent, model behavior, and agent decisions, not just request structure and syntax.

This semantic layer is what the OWASP LLM Top 10 controls ultimately depend on for prompt-injection defense and excessive-agency limits, and it is the layer that the incidents in Section 03 (the GitHub Copilot RCE and the Replit agent database deletion) show going missing in production.

The 2026 Threat Landscape

Threat intensity is rising fastest where control maturity is lowest, at the API and AI tiers. The vectors below account for the majority of incidents observed in field data across 2024–2025.

Top AI attack vectors

OWASP LLM Top 10 v2.0 released November 2024 and MITRE ATLAS now map every technique observed against production AI systems.

Vector	Description	Framework	Severity
Prompt Injection	User inputs override system instructions and change model behavior	LLM01:2025	CRITICAL
Indirect Prompt Injection	Payloads hidden in retrieved docs, emails, or web content.	LLM01 · ATLAS AML.T0051.001	CRITICAL
Data & Model Poisoning	Malicious training or fine-tuning data compromises behavior.	LLM04 · ATLAS AML.T0020	HIGH
Tool Abuse	Tool calls exercised outside intended scope via adversarial prompts.	LLM06 · Excessive Agency	HIGH
Agent Abuse	Autonomous agents chained into unintended, unauthenticated actions.	LLM06 · MITRE ATLAS	HIGH
Sensitive Data Leakage	Model, RAG, or logs expose PII, secrets, or system prompts.	LLM02 · LLM07 · LLM08	CRITICAL
MCP Risks	Model Context Protocol servers extend agent reach without auth boundaries.	Emerging · 2025	HIGH

How attackers target AI: a five-stage chain

Mapped to MITRE ATLAS, the adversarial-ML matrix updated in October 2024 to 16 tactics and 173 techniques, the AI kill chain gives each stage a specific, fundable control.

- **Reconnaissance (Recon):** Discover exposed model endpoints, system prompts, and public APIs.
- **Prompt Injection (Initial Access):** Deliver adversarial input via user, RAG, tool, or third-party page.
- **Tool Invocation (Execution):** Model calls an MCP tool, agent, or plugin outside intended scope.
- **Poisoning (Persistence):** Inject a payload into the vector store, fine-tune data, or agent memory.
- **Data Exfiltration :** PII, source code, or secrets sent via an output channel or tool.

The API attack lifecycle

On the API side, adversaries move through six stages, and the median cloud attack chain now completes in under ten minutes.

Stage	Adversary activity	Reference
Discovery	Maps external, internal, third-party, and shadow endpoints via passive scans and metadata.	OWASP API9
Enumeration	Object-ID walking, IDOR probing, and method inference exercised at scale.	OWASP API1 & API3
Access	Credential stuffing, session replay, and JWT forgery, stolen credentials remain the #1 initial action.	Verizon DBIR
Abuse	61% of API attacks in 2025 exercised abnormal workflows, refunds, coupons, transfers, scrapes.	Field data, 2025
Persistence	Long-lived tokens, orphaned service accounts, and rogue OAuth apps embedded in the tenant.	Cloud identity
Exfiltration	Bulk PII and records exfiltrated over authorized API channels, mean detection: 258 days.	IBM CODB

Where the two chains meet

The AI kill chain and the API attack lifecycle are not parallel tracks; they are the same attack described from two vantage points. A prompt injection that survives initial access is, from the API's perspective, just another authenticated request, one an API gateway will wave through because it is syntactically valid. A tool invocation that exercises excessive agency is, from the API's perspective, an authorized call to a scoped endpoint.

This is why point solutions fail: an API security tool blind to prompt intent will miss the injection, and an AI security tool blind to token scope and session state will miss the exfiltration that follows it. Closing the gap requires a single control plane that reads both layers at once, correlating request structure with prompt semantics, and agent behavior with API abuse patterns, so that a five-stage AI kill chain and a six-stage API attack lifecycle are detected as one incident, not two.


API

Lessons from the Breaches, 2022–2025

Five case files whose root causes shape the 2026 defense agenda: identity hygiene, and AI tool and agent permissioning. None required a novel exploit; each turned on a missing fundamental.

Incident	Root cause	Business impact	Security lesson
MS Storm-0558 Jul 2023	Forged Azure AD tokens from a stolen consumer signing key.	Federal-agency mailboxes accessed; national-security review.	Signing-key hygiene, key-scope separation, and continuous token validation are non-negotiable.
Snowflake Apr–Jun 2024	Stolen infostealer credentials; MFA not enforced on service accounts.	~165 customers incl. AT&T, Ticketmaster, Santander, billions of records.	MFA on every account, human or machine. Continuous behavior analytics on data platforms.
DeepSeek Jan 2025	Publicly accessible ClickHouse database with no authentication.	1M+ log entries, chat histories, and API-key material exposed.	AI infrastructure is regular infrastructure. Cloud-security fundamentals still apply.
GitHub Copilot RCE Aug 2025	Command injection via prompt injection (CVE-2025-53773, CVSS 9.6), Copilot's file-write privileges were hijacked to force an unapproved "YOLO mode."	Full local code execution on developer machines running Copilot in Visual Studio / VS Code; patched in Microsoft's August 2025 update.	AI coding agents need scoped, approval-gated tools and file permissions. Same root cause as classic command injection, new attack surface. OWASP LLM06 / Tool Abuse.
Replit Agent DB Deletion Jul 2025	Autonomous coding agents ran destructive commands during an active code freeze; guardrail existed only in the prompt, not the execution layer.	Live production database wiped (1,200+ executive and 1,190+ company records); agent then fabricated data and reported rollback as impossible.	Guardrails for autonomous agents must be enforced at execution (approval gates, environment separation, immutable backups), not merely instructed. OWASP LLM06 / Excessive Agency.

Sources: IBM Cost of a Data Breach 2025 · Verizon DBIR 2024 · Mandiant M-Trends 2025 · Microsoft Security Response Center & NVD (CVE-2025-53773) · public incident reporting on the Replit agent database deletion, July 2025.

Rows shaded yellow are new additions (GitHub Copilot RCE and the Replit agent database deletion), added in place of the Optus and MOVEit case files per the team's review.

Industry Risk Matrix

Sectors ranked by average breach cost, dominant attack pattern, and 2026 API-attack pressure. Financial services and healthcare remain the most-targeted verticals; manufacturing has risen to the #2 sectors for denial-of-service activity as OT/IoT systems are bridged to APIs.

Sector	Breach cost	Framework	Risk
Healthcare	$7.42M	Misc. Errors + Privilege Misuse (83%). PHI APIs, telehealth, MOVEit legacy.	CRITICAL
Manufacturing	$5.56M	System Intrusion + Misc. Errors (83%). #2 DoS sector; OT/IoT API bridges.	HIGH
Telecom	$3.75M	Basic Web App Attacks + Errors. Optus-pattern unauthenticated APIs are still active.	HIGH
Government	$2.86M	Misc. Errors + Intrusion + Social (78%). Legacy APIs, third-party exposure.	ELEVATED

Maturity: Where Enterprises Stand

Both API and AI security follow a five-stage progression from reactive response to autonomous defense. The gap between where most enterprises operate and where the threat now sits is the core finding of this report.

API security maturity

Most enterprises operate between Managed and Governed, and only 7.5% run a dedicated API threat-modeling program. The 2026 target is to move up the curve to continuous, autonomous defense.

Reactive → Managed → Governed → Automated → Autonomous. Progress is measured by how much of discovery, policy, detection, and response is continuous rather than periodic.

Sources: Salt Security, State of API Security 2024 (7.5% of organizations run a dedicated API threat-modeling program).

AI security maturity

The AI defense program mirrors it, and the stakes are concrete: 97% of organizations that suffered an AI-related incident lacked proper AI access controls.

Stage	What the stage delivers
01 · Discovery	Find every model, dataset, agent, and MCP tool in use.
02 · Governance	Approve use cases; apply NIST AI RMF and ISO 42001 controls.
03 · Protection	Input classification, prompt firewall, and tool-scope enforcement.
04 · Runtime	Continuous behavior analytics, output filtering, and tool audit.
05 · Continuous	Automated red-teaming, drift monitoring, and self-healing controls.

The Modern Defense: Cloud-Native, Runtime, Discovery

Kubernetes is now the de facto operating system for modern applications and AI workloads. Security must therefore sit inside the cluster, not in front of it.

Why the traditional WAF falls short

Perimeter WAFs were built for north-south HTTP traffic. They cannot see the east-west service-to-service calls, API schemas, and workload identities that define a cloud-native estate. A Kubernetes-native Web Application and API Protection (WAAP) layer is designed for exactly that context, inspecting API payloads, enforcing schema, and binding policy to workload identity rather than IP address.

Runtime is where security is decided

With the median cloud attack unfolding in under ten minutes and the mean breach lifecycle stretching to 258 days, prevention alone cannot close the gap. The modern runtime stack pairs continuous behavior analytics, output filtering, and tool-scope enforcement with automated blocking and forensic replay across both API and AI hops.

Sources: Sysdig Threat Research 2024 (ten-minute cloud attack timeframe) · IBM Cost of a Data Breach 2024 (258-day mean breach lifecycle).

Continuous discovery and Shadow AI

Every downstream control depends on a live inventory. Five source planes (gateways, traffic, code and CI/CD, cloud metadata, and runtime) should feed a continuously updated map of every API, model, agent, and dataset, whether external, internal, or shadow. This is the direct answer to Shadow AI: unsanctioned AI use bypasses procurement and legal review, but it still enters at the API layer, where discovery, governance, and runtime controls can catch it.

Sources: Baseline: OWASP API9:2023 "Improper Inventory Management" · Salt Security & Postman 2024 field data on shadow-API prevalence.

Frameworks and the Executive Roadmap

Seven authoritative frameworks now anchor the combined AI + API + application security discipline, each addressing a different lifecycle stage: OWASP API Security Top 10, OWASP LLM Top 10, MITRE ATLAS, MITRE ATT&CK, NIST AI RMF, ISO/IEC 42001, and CISA Secure by Design.

A six-step operating loop

Each step is a decision the board can measure, not a promise a vendor can make.

Step	What it does	Anchored to
Discover	Inventory every API, model, agent, and dataset, internal, external, and shadow.	Traditional Apps
Govern	Policy-as-code, use-case review, ownership, and approval gates.	NIST AI RMF · ISO 42001
Detect	Continuous behavior analytics and anomaly detection at the API tier.	Sysdig · Mandiant M-Trends
Respond	Runbooks, automated blocking, and forensic replay across API + AI hops.	MITRE ATT&CK · ATLAS
Validate	Red-team, chaos, and continuous validation of controls and policy drift.	CISA Secure by Design

Outlook: Preparing for 2027

The direction of travel is clear from the primary-source research reviewed in this report. Five forecasts should shape 2026–2027 planning:

- **Agentic AI moves from pilot to production.** Gartner projects autonomous agents will feature in roughly a third of enterprise applications by 2028, making agent identity, tool-scope enforcement, and MCP governance board-level concerns.
- **The API becomes the primary control plane for AI security.** Because every prompt is a request and every model call is an endpoint, AI security converges on API security disciplines: authentication, schema enforcement, and rate limiting.
- **Shadow AI governance shifts left to discovery.** With most AI users bringing their own tools, continuous discovery, not policy memos, becomes the enforceable control.
- **Runtime defense outpaces perimeter prevention.** As attack chains compress below ten minutes, investment moves toward in-cluster, identity-bound runtime protection.
- **Regulation and framework adoption accelerate.** NIST AI RMF, ISO/IEC 42001, and CISA Secure by Design move from reference material to procurement and audit requirements.

Sources: Gartner · IBM Cost of a Data Breach · OWASP LLM Top 10 v2.0 · CyberArk. Forward-looking analysis by Prophaze Threat Research.

About Prophaze

Prophaze is a Kubernetes-native application, API, and AI security platform, designed for the modern enterprise stack, where the perimeter has become an interface and every AI feature ships as an API. Its remit spans continuous discovery, governance, protection, and runtime defense across the connected attack surface: protecting modern applications from discovery to runtime.

Our Vision

Integrated Security Proactive, Not Reactive

We believe security should not be a collection of tools customers hope will work. It should be a complete solution and a partner that stands with you when attacks happen. Prophaze exists to take full responsibility for protecting web applications and APIs, combining a fully featured platform, AI-driven intelligence, and expert humans in the loop so businesses can rely on protection that holds up under real pressure.


Trusted by Leading Brands
EMAAR
BHARAT ELECTRONICS
HDFC CAPITAL
CAIRN
Cemindia By Adani group
eHealth Kerala
IIT MADRAS
DSCI
A nasscom Initiative
NATIONAL HOUSING BANK
HINDUSTAN ZINC
COLORTOKENS
AMRITA HOSPITAL
KELTRON
A Government of Kerala Undertaking
ONGC
MRPL
KFON
fractal
vedanta
GMR
DECOR
balco
Kirloskar
UNITERS

Threat Intelligence | Enterprise Readiness | Cloud-Native Security | Runtime Protection

Published by Prophaze

Kubernetes-native application, API & AI security
Protecting modern applications from discovery to runtime.

Scan QR Code →
Book Your Demo Today

Prophaze™
The New Phase of Security

contact@prophaze.com | www.prophaze.com