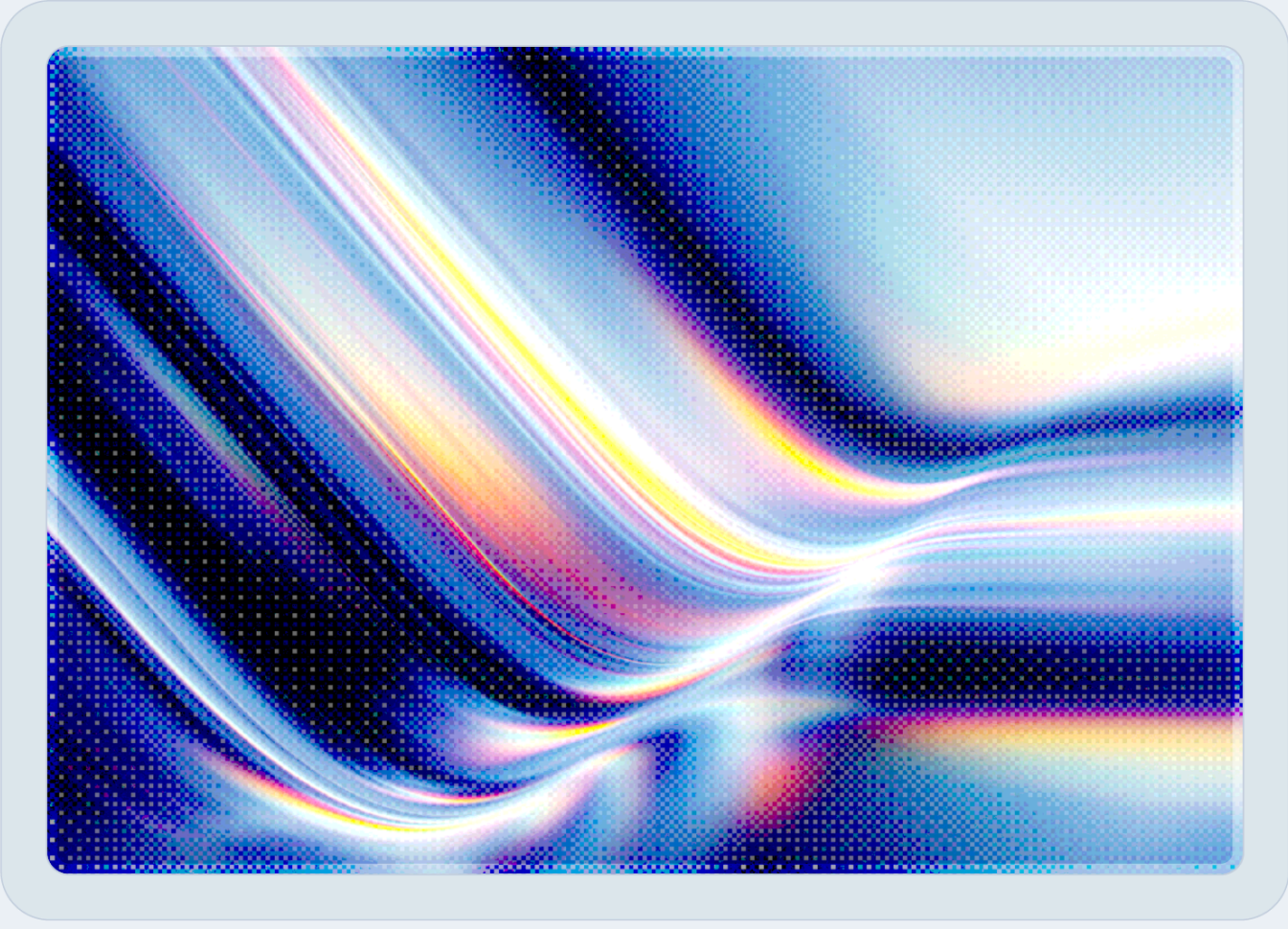
CSAI Foundation | Cloud Security Alliance

VulnOps: Vulnerability Management in the Age of AI

Why the Old Operating Model Is Failing and What Comes Next

2026-05-04

AI-assisted Rapid Research



CSAI

CSA cloud security alliance®

Key Takeaways

The volume, velocity, and character of software vulnerability disclosure have shifted faster than the operating model most organizations use to respond to them. Autonomous AI research systems are now discovering thousands of previously unknown defects in widely deployed code, the public infrastructure used to enrich and score those findings has formally conceded that it cannot keep pace, and the time between disclosure and active exploitation continues to compress. Traditional vulnerability management — built around periodic scans, severity-ranked queues, and quarterly patch cycles — was designed for a different threat landscape. The discipline that replaces it is increasingly being called *VulnOps*, and while the term is not new, its operational meaning is changing under the pressures described in this note.

- AI-driven discovery is producing CVE submission rates that exceed the structural capacity of the National Vulnerability Database, which on April 15, 2026 announced a risk-based triage policy that limits full enrichment to vulnerabilities listed in CISA's Known Exploited Vulnerabilities catalog, software designated under Executive Order 14028, or software in active federal use, leaving the majority of new CVEs without full enrichment [1].
- The traditional vulnerability management operating model — scan, ticket, prioritize by CVSS, patch on a calendar — was designed against a backdrop of roughly 10,000 CVEs per year. It is now being asked to absorb four to five times that volume [2], with no comparable scaling of public enrichment infrastructure visible in NVD's published throughput [3].
- VulnOps reframes vulnerability management as a continuous flow problem rather than a queue problem, drawing on the operational vocabulary of DevOps, SRE, and incident response to make the practice tractable at modern scale.
- The discipline must work asymmetrically. The same principles apply to a Fortune 100 enterprise and a county government below the security poverty line, but the implementation patterns diverge sharply, and a guidance set that ignores either context will fail in practice.

Background: The Operating Model Is Breaking

Vulnerability management as an enterprise practice was codified in the early 2000s around three assumptions that no longer hold. The first was that the rate of new vulnerability disclosure was approximately predictable, growing roughly with the size of the deployed software base. The second was

that public infrastructure — the CVE program, the National Vulnerability Database, vendor advisories — could be relied upon to enrich raw disclosures with the metadata needed to act on them. The third was that the time between public disclosure and weaponized exploitation provided a defensible window, typically measured in weeks or months, during which a competent program could plan, test, and deploy a remediation. Each of those assumptions has now eroded materially.

CVE submission volume has roughly tripled since 2020, and the inbound pace in early 2026 has continued to climb relative to 2025 [3]. The acceleration is no longer linear and no longer attributable to a single cause. Application complexity has continued to grow, the CVE Numbering Authority program has continued to expand, and a new class of autonomous discovery systems has begun producing findings at throughputs no manual program can match. Google's Big Sleep agent, in an August 2025 disclosure, reported having surfaced more than twenty previously unknown vulnerabilities in widely used open-source projects since its launch the previous fall [4]. DARPA's AI Cyber Challenge demonstrated that autonomous systems could discover and patch the majority of injected synthetic vulnerabilities in critical open-source projects — including the Linux kernel, SQLite, and cURL — and surfaced eighteen previously unknown real vulnerabilities in the process [5]. Other frontier AI laboratories have begun publicly describing internal programs that produce zero-day candidates at throughputs no prior research effort has matched. Whether or not any single claim is reproduced at the stated scale, the floor of plausible AI-driven discovery output is already well above what manual triage was built to absorb.

Public tracking infrastructure has not scaled with that surge. NIST's NVD enriched 41,925 CVEs in 2025 — a forty-five percent year-over-year increase and the largest output in the database's history — and still finished the year with a backlog covering most of the inbound queue [3]. On April 15, 2026, NIST formally acknowledged the structural mismatch and announced a risk-based triage policy that limits full enrichment to vulnerabilities listed in CISA's Known Exploited Vulnerabilities catalog, software designated under Executive Order 14028, or software in active federal use [1]. The policy is realistic given the constraints, but its consequence is that the majority of new CVEs will reach enterprise security tooling without the CVSS scores, CPE identifiers, and CWE classifications that automated scanners and patch management systems depend upon. The output of that pipeline is now structurally partial as a matter of declared policy.

The third assumption — that disclosure provides a defensible exploitation window — has eroded most visibly. Multiple incident response and threat intelligence programs report median time-to-exploitation measured in days for high-impact vulnerabilities, with some weaponizations now arriving within hours of public disclosure [6]. The compression is driven partly by the same automation that enables AI-augmented discovery: an attacker who can ingest a public advisory, generate a working proof of concept, and deploy at scale across an internet-facing target list operates on a clock that a calendar-based remediation cadence cannot meet on its own, absent compensating controls.

What VulnOps Is — and What It Isn't

VulnOps is best understood as an operating discipline rather than a product category. The label has been circulating in security operations communities for several years and predates this research note; CSA's contribution is not to coin the term but to articulate what the discipline must contain to be useful in 2026. The closest analogues are DevOps and Site Reliability Engineering, which similarly emerged as informal labels for a set of cultural and operational practices before crystallizing into a recognizable discipline with its own patterns, tooling, and career paths. VulnOps occupies the same conceptual space for the work that traditionally fell under "vulnerability management."

A working definition: VulnOps is the continuous, flow-oriented operational discipline of identifying, prioritizing, remediating, and verifying software defects across the systems an organization owns or depends upon, with first-class treatment of the supply chain, the software development lifecycle, and the operational telemetry that surrounds both. It treats vulnerability response as a stream of small, frequent, partially automated actions rather than a periodic batch process. It organizes work around the question "what is the system trying to tell us, and what should change in response?" rather than "which tickets are open this quarter?"

The shift in framing is the substantive point. A queue-based model implicitly assumes that throughput is bounded by the size of the team draining the queue, and that the queue itself is approximately stable in shape. A flow-based model assumes that inbound rate is the dominant constraint, that work-in-progress is a measurable signal, and that the system must be tuned for cycle time rather than depth. The instrumentation looks different: a VulnOps program measures lead time from disclosure to verified remediation, change-failure rate on patches, recurrence rate of previously remediated findings, and saturation of the remediation pipeline against inbound discovery rate. Those are SRE metrics applied to vulnerability work.

VulnOps is not a replacement for vulnerability scanners, software composition analysis tools, or patch management platforms. Those products remain part of the underlying instrumentation. What VulnOps reorganizes is how their outputs are consumed, prioritized, and acted upon. A scanner that produces 200,000 findings against a deployed estate is not the problem; the problem is the absence of a flow-shaped operating model that can convert that telemetry into a survivable cadence of action.

Forces Reshaping the Practice

Four forces are pushing organizations toward an explicit VulnOps posture, and any serious guidance must address each.

The first is the reshaping of the discovery side of the lifecycle by AI-augmented research. Autonomous and semi-autonomous systems have demonstrated, repeatedly, the ability to surface non-trivial defects in heavily reviewed open-source projects, in commercial operating systems, and in browser engines. A plausible economic implication, supported anecdotally by reductions in bug-bounty time-to-find for participating teams, is that the marginal cost of finding a previously unknown vulnerability has fallen sharply, both for defenders and for offensively motivated researchers. The defensive consequence is that the volume of "known" issues facing any given enterprise will keep rising independent of the enterprise's own actions.

The second is the partial failure of public enrichment infrastructure. The NVD policy change is the most prominent example, but it is not isolated. The CVE Program has weathered repeated funding scares, MITRE's role has been the subject of ongoing federal contract negotiations, and a multi-source ecosystem of alternative databases — VulnCheck, ENISA's EUVD, and vendor-specific advisories — has emerged in the gap. Programs that built their pipelines on a single authoritative reference will need to integrate multiple sources and reconcile their outputs, which is a non-trivial engineering effort.

The third is the compression of the exploitation window, discussed above. The operational implication is that any program whose remediation cadence is measured in calendar weeks for internet-exposed systems is operating outside the threat model. That requires not only faster patching but a different decomposition of the work: continuous discovery on the asset side, automated triage that does not depend on enrichment that may never arrive, and remediation paths that are pre-built rather than designed at the moment of need.

The fourth is the emergence of agentic AI itself as both a defender's tool and an attacker's surface. AI agents are increasingly embedded in the development pipeline, in production systems, and in the security stack. They introduce new defect classes — prompt injection chains, tool-poisoning attacks against agent integrations, model-context-protocol command execution flaws — that do not map cleanly onto traditional CVE taxonomies but are nevertheless vulnerabilities in the operational sense [7]. A VulnOps practice that ignores agentic AI surfaces is likely to be measuring the wrong estate within a small number of years, particularly as agent integration in production systems accelerates.

Core Principles

A VulnOps practice that takes the four forces above seriously tends to organize around a small set of principles.

Continuous discovery replaces periodic scanning. The unit of work is not a quarterly scan report but an event stream from agent-based asset discovery, software composition analysis on every build, attack-surface monitoring against external-facing assets, and ingestion of multiple vulnerability databases rather than a single authoritative source. The instrumentation runs continuously and produces signal continuously.

Risk-based prioritization replaces severity ranking. CVSS base scores remain useful as a coarse filter, but they are increasingly insufficient as a primary prioritization signal because a growing share of CVEs will arrive without one. Programs are moving toward composite signals that combine exploitability data such as EPSS and CISA KEV, asset context drawn from configuration management and identity systems, and threat intelligence about active campaigns. The composite is more expensive to maintain than a CVSS-only model but is the only approach that survives a partially enriched CVE feed.

Automated remediation flow replaces ticket queues for the high-volume tail. The principle is that the response to a vulnerability in a containerized service should default to an automated rebuild against a patched base image, with a tested rollback path, rather than a human ticket. Tickets remain appropriate for the long tail of legacy systems and exception cases, but they should not be the default disposition for the majority of findings.

Software supply chain awareness is treated as a first-class concern. CSA's position, consistent with industry frameworks for software supply-chain transparency such as SLSA and the practices set out in NIST SP 800-218, is that a vulnerability management program should be able to answer "where is this dependency in our stack?" within minutes. Software bill of materials data, dependency-graph indexing, and visibility into transitive dependencies are infrastructure under that standard, not optional extras.

The discipline is implemented asymmetrically across organizational scales. The same principles apply to a global bank and to a regional hospital system, but the implementation patterns differ significantly. The guidance must address both large enterprises and organizations operating below the security poverty line, a concept long associated with practitioner Wendy Nather and arguably more relevant than at any point in its fifteen-year history [8]. A VulnOps reference that only describes the well-resourced case is incomplete.

Finally, the practice is AI-aware in both directions. AI is integrated into the operational stack — for triage, for code-fix generation, for natural-language summarization of advisories — and AI-driven systems are themselves part of the protected estate. Both halves of that integration — AI as an operational tool and AI systems as a protected surface — are essential to a complete VulnOps practice.

Recommendations

For organizations beginning the transition, three concrete starting points apply broadly. The first is to instrument the inbound rate of vulnerability findings against the throughput of remediation, treat the ratio as a saturation metric, and report it to security leadership monthly. A program that is consistently saturated cannot be fixed with more tickets, and the metric makes the structural problem visible.

The second is to diversify vulnerability intelligence sources rather than depending on a single enriched feed. The NVD remains useful for the subset of CVEs it now prioritizes, but it should be paired with at least one alternative source covering the unenriched tail and with exploitability signals such as the CISA Known Exploited Vulnerabilities catalog and EPSS [9][10]. The integration cost is real, but for programs whose decisions depend on the unenriched tail it is increasingly justified by the operational cost of acting on incomplete data.

The third is to begin treating remediation as code. The default response to a vulnerable container image, package version, or configuration setting should be a tested, repeatable change committed to a repository and deployed through the same pipeline as any other change. The papers on CI/CD and AI agent integration develop this in detail, but the cultural shift — that remediation is engineering work, not a service-desk ticket — is the foundational move.

References

[1] National Institute of Standards and Technology. "NIST Updates NVD Operations to Address Record CVE Growth." NIST News, April 15, 2026.

[2] CVE Program. "CVE Metrics." cve.org.

[3] National Institute of Standards and Technology. "NVD Dashboard." NVD, accessed May 2026.

[4] Google. "Cybersecurity Updates: Summer 2025 — Big Sleep AI Agent Discovers Real-World Vulnerabilities." Google Blog, August 2025.

[5] Defense Advanced Research Projects Agency. "AI Cyber Challenge." DARPA AIxCC, 2025.

[6] Mandiant. "M-Trends 2025 Special Report." Google Cloud, 2025.

[7] Cloud Security Alliance Labs. "MCP by Design: RCE Across the AI Agent Ecosystem." CSA Labs, April 20, 2026.

[8] Nather, Wendy. "The Security Poverty Line, and Junk Food." Idoneous Security (personal blog), December 2011.

[9] Cybersecurity and Infrastructure Security Agency. "Known Exploited Vulnerabilities Catalog." CISA.

[10] FIRST. "Exploit Prediction Scoring System (EPSS)." Forum of Incident Response and Security Teams.

