Derailed

2026 Application Security Benchmark Report

OX Security 2nd Annual Report

Analyzing 216 Million Application Security Findings

250 Organizations | 90-Day Analysis Period

Table of Contents

Executive Summary

This Train is Officially off the Track

The second annual OX Application Security Benchmark Report documents a measurable and significant escalation in the application security challenge facing organizations worldwide. Analyzing over 216 million security findings across 250 organizations — more than double last year's dataset — the data presents a clear picture: the volume, severity, and complexity of application security risk are all increasing.

The total number of publicly disclosed CVEs reached 48,185 in 2025, up from 40,008 in 2024 — a 20% year-over-year increase, continuing a decade-long acceleration. This growth in the raw vulnerability landscape is compounded by the widespread adoption of AI-assisted development, which has dramatically increased code output and, with it, the rate at which new security issues are introduced into software pipelines.

The average organization now faces 865,398 security alerts — a 52% increase from 569,354 in 2025. And the data suggests this trajectory will not plateau.

Prioritization continues to be the most effective mechanism for making this volume manageable — though the data raises legitimate questions about how long that will remain true in an exponentially escalating landscape.

After applying prioritization methodology — consolidating findings, then analyzing exploitability, reachability, and impact against each organization's unique architecture — the average organization's alert load is reduced to 8,081 actionable issues, of which only 795 are critical. What has changed materially from last year is the composition of those critical issues: the absolute count of critical findings after prioritization has risen nearly 4x, from 202 to 795 per organization, increasing critical workloads significantly.

Key figures at a glance:

216M+
Findings Analysed
vs 101M in 2025

250
Organizations
vs 178 in 2025

865K
Avg Alerts/Org
+52% YoY

795
Critical/Org (post-prio)
vs 202 in 2025

Executive Summary

Key figures at a glance

Introduction

AI as the Driving Force of Risk (and Security)

Last year's report introduced the concept of the alert fatigue crisis — a structural mismatch between the volume of security findings generated by modern tooling and the capacity of security and development teams to act on them. This year's data confirms that the underlying dynamics driving that crisis have not stabilized; they have intensified.

Two forces are compounding each other. First, the raw vulnerability landscape continues to expand at pace. Second, and more significantly for this year's findings, AI-assisted development has entered mainstream software engineering practice at scale. The widespread adoption of AI code generation tools has increased development velocity substantially — but it has also introduced a new category of structural risk: AI-generated code that reaches production pipelines without the same security review cadence that human-authored code typically undergoes.

The OX Research Team's "Army of Juniors" report, published in 2025, documented this dynamic in detail: AI coding assistants tend to replicate common vulnerability patterns — in scale and velocity. New development cycles now produce code that passes functional review but carries security weaknesses. The benchmark data in this report reflects the downstream effects of that trend at the organizational level.

CVE disclosures reached 48,185 in 2025 — up 20% from 40,008 in 2024, continuing a decade of unbroken growth.

The implications extend beyond raw numbers. As the volume of findings grows, it is clear that human-based security is already obsolete — no organization will survive the development shift without an adaptation of its security operations.

This report's second annual dataset offers, for the first time, a year-over-year view of how these dynamics manifest at the organizational level, and across specific sub-industries. The methodology has also evolved: where last year's benchmark reflected a 90-day aggregated view, the 2026 industry benchmark data reflects a daily dashboard methodology, capturing a more operationally accurate picture of what security teams are managing on any given day.

The widespread adoption of AI code generation tools has increased development velocity substantially — but it has also introduced a new category of structural risk.

Methodology

The 2026 Benchmark Report is based on an analysis of 216 million application security findings collected from 250 organizations over a 90-day period (Q4 2025). The OX Research Team employed a three-step methodology consistent with last year's approach, with one significant update to the industry benchmark component.

Data Collection and Consolidation

Security findings were aggregated and consolidated from organizations' existing security tools — including SAST, secrets detection, SCA, and others — alongside open-source intelligence feeds and vulnerability databases.

Data Enrichment

Each security event was enriched with new contextual data points, including mapping cloud assets back to their code origin using OX's proprietary technology, analyzing component dependencies, and assessing the business criticality of affected systems.

Context Analysis and Prioritization

Prioritization was applied using multiple risk labels per issue to assess real-world impact and actual risk level. This incorporated factors including exploitability, reachability, business impact, and environmental context.

Methodology Update: Industry Benchmark

The 2025 report's industry benchmark reflected cumulative findings over a 90-day period, segmented by broad industry vertical and company size. For the 2026 edition, the benchmark has been updated to reflect a daily dashboard view — capturing the volume of findings and critical alerts that organizations are managing at any given point in time. This change produces figures more directly comparable to the operational reality that security teams face, and replaces the enterprise/SMB segmentation with a sub-industry breakdown that provides greater specificity.

Study Parameter	2025	2026
Findings analyzed	101,344,969	216,349,575
Organizations	178	250
Duration	90 days (Q4 2024)	90 days (Q4 2025)
Benchmark methodology	90 days aggregate, Enterprise/SMB	Daily dashboard, Sub-industry

Research

The Threat Landscape

The number of publicly disclosed application security vulnerabilities has grown continuously for over a decade. In 2015, CVEdetail recorded 6,494 vulnerabilities. By 2024, that figure had reached 40,008. In 2025, it climbed further to 48,185 — a 20% single-year increase, and a near-8x increase over the past decade.

48,185 CVEs disclosed in 2025 | 40,008 in 2024 | +20% year-over-year

FIRST (Forum of Incident Response and Security Teams) projects that this trend will continue, with an estimated 41,000—50,000 new weaknesses expected in 2026. The structural causes driving this growth — increasing software complexity, the proliferation of open-source dependencies, and now the acceleration of AI-assisted code generation — show no signs of abating.

AI-assisted development deserves particular attention in this context. The OX "Army of Juniors" report documented a consistent pattern: AI coding assistants, trained on large corpora of existing code, tend to reproduce common vulnerability patterns alongside functional code. The report also identified 10 anti-patterns — recurring practices that run counter to security best practice. Critically, while vulnerability density in AI-generated code is not necessarily higher than in human-written code, the sheer velocity of AI-assisted development introduces 2x, 5x, or 10x more lines of code per day. This velocity effect directly explains the 52% increase in average total issues per organization documented in this report.

The signal-to-noise ratio deteriorates, and the risk of critical issues going unaddressed — lost in a backlog of lower-priority findings — increases proportionally.

While vulnerability density in AI-generated code is not necessarily higher than in human-written code, the sheer velocity of AI-assisted development explains the 52% increase in ave. total issues per organization documented in this report.

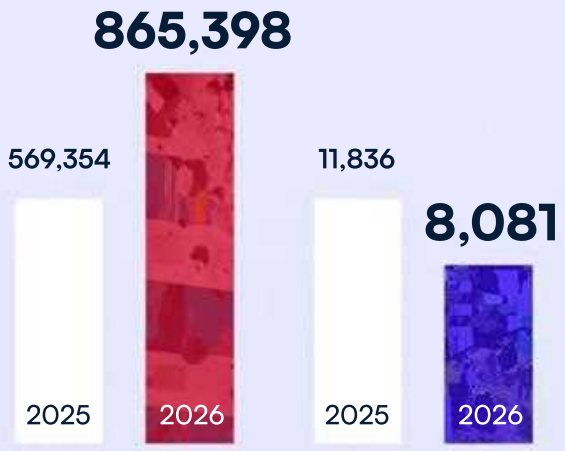
Research

Key Findings: Year-Over-Year Escalation

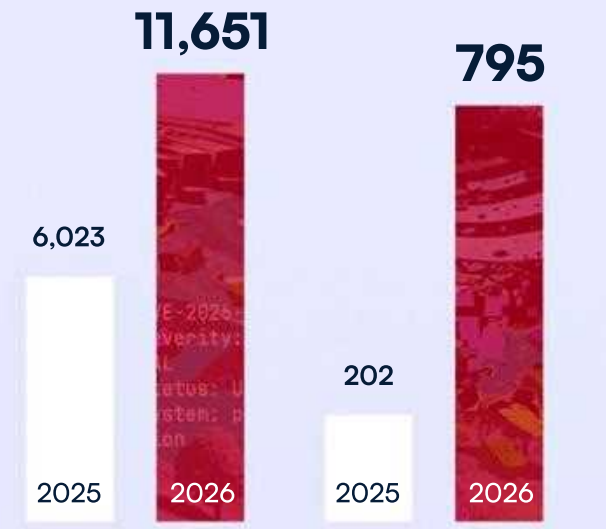
1. Alert Volume and Organizational Impact

The most immediate story in this year's data is the scale of growth. Organizations participating in this benchmark saw their average raw alert volume increase by 52% year-over-year, from 569,354 to 865,398 findings, suggesting that AI-assisted development is amplifying the velocity at which new vulnerabilities enter organizational codebases.

Total Issues Count



Critical Issues Count



Prioritization continues to reduce the overall alert load substantially — from 865,398 to 8,081 actionable issues, of which only 795 are critical. However, the post-prioritization critical count has risen from 202 to 795 per organization.

This is a meaningful shift: the number of issues that genuinely require immediate security attention has nearly quadrupled year-over-year. Security teams are not just dealing with more noise — they are dealing with more genuine risk and critical workloads.

2. The Critical Issue Ratio

The proportion of raw alerts that are classified as critical after prioritization has increased year-over-year. In 2025, 202 critical issues out of 569,354 raw findings represented approximately 0.035% of the total. In 2026, 795 out of 865,398 represents approximately 0.092% — a near-tripling of the critical ratio.

This trend indicates that while the majority of alerts remain non-critical and can be systematically deprioritized, the genuine risk embedded within the overall alert volume is growing at a faster rate than the volume itself. The implication for triage strategy is significant: the critical-to-noise ratio is deteriorating, making prioritization more important, not less, as the landscape evolves.

The critical issue ratio has nearly tripled year-over-year: from ~0.035% to ~0.092% of raw findings.

3. Severity Factors

Not all alerts carry equal weight — and the data reveals exactly what separates the critical minority from the deprioritized majority.

OX's prioritization methodology applies multiple risk labels per issue, drawn from three dimensions: exploitability, damage potential, and reachability. The following breakdown reflects which factors most frequently drove issues up or down the priority stack across 216 million findings.

Four of the five top risk-elevating factors are damage-related rather than exploit-related.

Factors That Increase Risk

Factor Type	Factor Name	% of All Issues
Impact	High Business Priority	**27.76%**
Impact	PII Processing	**22.08%**
Impact	CVSS - High Severity	**20.55%**
Impact	Critical Business Priority	**19.49%**
Exploitability	Multiple Public Exploits	**17.88%**

The most striking pattern in the risk-increasing factors is that **business context dominates over technical severity**. Four of the five top risk-elevating factors are damage-related rather than exploit-related.

The most frequently applied risk-elevating label is **High Business Priority**, appearing in 27.76% of all findings — meaning more than one in four issues flagged across the dataset touched a system considered business-critical. **Critical Business Priority** (19.49%) follows closely.

PII Processing appeared in 22.08% of findings, reflecting how often sensitive data handling is in the blast radius of potential vulnerabilities. **CVSS High Severity** appeared in 20.55% of findings, while **Multiple Public Exploits** — the only purely technical exploitability signal in the top five — appeared in 17.88% of findings.

Factors That Reduce Risk

Factor Type	Factor Name	% of All Issues
Exploitability	EPSS - Low Exploit Risk	**36.27%**
Exploitability	Public Exploit Unavailable	**24.47%**
Impact	Low Business Priority	**14.26%**
Impact	Medium Business Priority	**12.71%**
Reachability	Development Dependency	**11.26%**

On the deprioritization side, exploit probability dominates.

EPSS Low Exploit Risk alone accounts for 36.27% of all findings analyzed — meaning that for more than a third of all alerts, the statistical likelihood of exploitation is low enough to warrant removing them from the active workload.

Public Exploit Unavailable adds another 24.47%, reinforcing the same principle: without a practical path to exploitation, a vulnerability's theoretical severity carries limited operational weight. Together, these two exploit-related factors are the most prevalent deprioritization signals in the dataset — each appearing independently across tens of millions of findings.

Low and Medium Business Priority labels appeared in 14.26% and 12.71% of findings respectively, and **Development Dependency** in 11.26% — confirming that vulnerabilities in code that never reaches production remain a persistent source of noise.

The data makes a consistent case: what a vulnerability touches matters more than what it scores. Business context — the criticality of the affected system, the data it processes, and its role in the organization — is the primary lens through which risk should be assessed.

Research

Daily Dashboard

Industry Average Breakdown

Sub-Industry	Total Raw	→	Total Prioritized	Critical Raw	→	Critical Prioritized	Critical Signal %
Insurance	68,104	→	10,925	5,735	→	1,197	1.76%
Gaming	43,648	→	5,101	1,856	→	548	1.26%
Media	113,305	→	13,938	6,343	→	1,335	1.18%
Automotive	253,438	→	36,973	18,285	→	2,722	1.07%
Financial Services	187,886	→	24,459	11,836	→	1,993	1.06%
Healthcare	123,033	→	10,552	6,852	→	1,268	1.03%
Technology	173,754	→	15,510	7,125	→	1,485	0.85%
Software Development	197,464	→	14,477	7,781	→	1,527	0.77%
Biotechnology	126,253	→	5,628	1,985	→	573	0.45%

Sub-Industry Noise Rate			
Insurance	98.24%	Healthcare	98.97%
Gaming	98.74%	Technology	99.15%
Media	98.82%	Software Development	99.23%
Automotive	98.93%	Biotechnology	99.55%
Financial Services	98.94%		

Research

Industry Benchmark: Sub-Industry Analysis

This year's benchmark introduces a sub-industry lens, replacing last year's broad vertical and size-based segmentation. The data reflects a daily dashboard view — findings and critical alerts as they exist on any given operational day — rather than a cumulative 90-day aggregate. This methodology change provides a more accurate representation of the operational burden that security teams face.

Several patterns emerge from this sub-industry data:

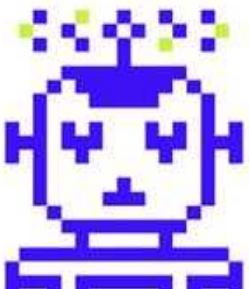
Insurance leads on critical signal percentage. At 1.76%, Insurance organizations have the highest proportion of findings that survive contextual prioritization as genuinely critical. This reflects both the sensitivity of the data these organizations manage and the regulatory scrutiny they operate under.

Automotive shows the highest absolute volumes. With 253,438 raw findings — the highest of any sub-industry — Automotive organizations are managing an increasingly complex security surface, driven by the convergence of software-defined vehicle systems, connected infrastructure, and supply chain complexity.

Software Development and Technology have the lowest critical signal percentages. At 0.77% and 0.85% respectively, these sectors show a relatively lower proportion of critical findings, despite high raw volumes. This may reflect greater security maturity and more developed tooling ecosystems in organizations that are themselves builders of software infrastructure.

Biotechnology is an outlier. At 0.45% critical signal and 5,628 total prioritized findings, Biotechnology organizations present a significantly lower critical ratio. This warrants monitoring in future editions: it may reflect genuine security maturity, a narrower application footprint, or differences in the types of vulnerabilities present in this sector's software stack.

Financial Services, while no longer the highest-volume sector in this breakdown, retains a 1.06% critical signal percentage — consistent with its historically elevated risk profile and the continued attractiveness of financial data to attackers.



Research

From Remediation Costs to Prevention Economics

The data in this report describes a system under pressure. But pressure has a cost — and that cost is worth making explicit.

When organizations lack the ability to systematically distinguish critical findings from non-critical ones, the downstream effects are compounding: security and development hours are consumed by low-priority triage; genuine critical issues take longer to reach remediation; and the organizational friction between security teams and development teams — a persistent feature of AppSec programs — intensifies.

The year-over-year growth suggests this cost trajectory is accelerating. **A 52% increase in raw alert volume, combined with a near-tripling of the critical issue ratio,** means that organizations face a materially worse operational situation in 2026 than they did in 2025.

One additional cost dimension worth noting is the point-in-pipeline effect: vulnerabilities identified and addressed early in development are substantially cheaper to remediate than those that reach production. **The traditional "shift left" model tried (and failed) to move security earlier in the development cycle for exactly this reason.** But AI-assisted development introduces a more fundamental question: if AI models are now generating the code that is driving the backlog explosion, the most logical intervention point is not after the code is written — it is at the moment of generation itself. Embedding security context into AI code generation models means vulnerabilities are prevented from being introduced rather than detected and remediated after the fact.

This distinction matters for the cost curve. **Detection and remediation — however efficient — are responses to vulnerabilities that already exist.** Prevention operates upstream of that entire cycle. Given the velocity at which AI-assisted development is scaling, prevention-led security may be the only approach capable of making the year-over-year growth manageable — and ultimately, of turning the curve around.

When organizations lack the ability to distinguish critical findings from non-critical ones, security and development hours are consumed by triage; critical issues take longer to reach remediation; and the friction between teams — a persistent feature of AppSec programs — intensifies.

Conclusions

On the Verge of Re-Boot

Three findings from this year's benchmark warrant particular attention as the industry looks ahead.

The volume problem is outpacing remediation capacity. A 52% increase in average alert volume year-over-year, against a backdrop of 20% CVE growth, suggests that AI-assisted development is becoming a material driver of security backlog. This is not a temporary effect: as AI coding tools become more deeply embedded in development workflows, the pipeline between code generation and security review will remain a primary area of pressure.

The critical ratio is rising, not falling. The near-tripling of the post-prioritization critical issue ratio indicates that genuine risk is growing within the overall alert volume — not just total noise. Organizations that treat this as a constant percentage are likely underestimating the challenge.

Sub-industry risk profiles differ in ways that matter. Insurance, Automotive, and Financial Services face the highest critical signal percentages. Software Development and Biotechnology sit at the other end of the spectrum.

The data points to a widening gap between the pace at which vulnerabilities are introduced and the capacity of teams to manage them. The trend line is clear — the question is whether organizational security practices and solutions will keep pace.

What is consistent across all segments is that context-aware prioritization — incorporating exploitability, reachability, and business impact — remains the primary mechanism for making the overall volume tractable. The 2025 finding that the vast majority of raw alerts can be safely deprioritized holds in 2026. What has changed is the scale of what remains after that prioritization, and the urgency attached to it.

The path most consistent with the data is one that addresses vulnerabilities as early in the development lifecycle as possible — before they accumulate into the volumes this benchmark reflects, before the attack surface becomes overgrown, and before attackers can exploit it.

About OX

OX Security is a leader in application security, providing comprehensive coverage across the entire software development lifecycle, from AI code generation to cloud runtime. OX centralizes security across the entire code journey, tracing every risk back to its source, so security teams can move from fragmented tooling and blind spots to complete, unified product security built for prevention.

www.ox.security

2026 Application Security Benchmark Report