Software Analyst®
Cyber Research

SEPTEMBER
2025

THE CONVERGENCE OF SIEMS AND DATA LAKES: MARKET EVOLUTION, KEY PLAYERS AND WHAT'S NEXT

- AQSA TAYLOR
- FRANCIS ODUM
- RAFAŁ KITAB

About Software Analyst Cybersecurity Research

SACR is a modern research and advisory firm built for today's cybersecurity leaders. We deliver in-depth, timely analysis across SOC operations, Identity, Network, Cloud, Application Security, Data, and AI Security; equipping CISOs, security teams, founders, investors, and practitioners with the insight they need to navigate high-stakes decisions.

With an engaged community of over **80,000** readers and followers, SACR connects with a global network of cybersecurity decision-makers and innovators. Our access to leaders across categories and industries gives us a direct line to the conversations shaping the market. By pairing these insights with rigorous technical analysis and continuous market tracking, we produce research that is both data-driven and grounded in the realities of modern security operations.

Whether you're seeking clarity on emerging technologies, evaluating vendors, or tracking market shifts, SACR delivers trusted, independent research designed to help you see clearly and decide with confidence.

We explore the newest frontiers of cybersecurity.

Whether you're looking at emerging vendors, evolving threats, or shifting architectures, or timely, opinionated insights help modern security leaders make smarter, faster decisions.

Table of Contents

Executive Summary

The SIEM market is undergoing one of its most significant shifts in decades, driven by the rise of security data lakes, pipelines, and advanced analytics platforms. What began as centralized log management has expanded into a battle over who controls the SOC's data layer—traditional SIEM vendors, cloud-native data platforms, or emerging security data pipeline providers. This report explores the convergence of SIEMs and data lakes, the competitive dynamics shaping the market, and the key players redefining how enterprises detect, analyze, and respond to threats.

We further explore how the SIEM is being modernized in 2025 to address legacy concerns, including rising costs, noisy data, and expanding SOC requirements. It serves as a field guide for CISOs, SOC leaders, and practitioners evaluating modern SIEM models. The report defines the core architecture patterns shaping the market, including pipeline-first designs, decoupled compute and storage with data lakes and federated query layers, and converged platforms that unify SIEM with UEBA, SDPP, XDR, SOAR, and exposure management. We've collaborated with a select number of large SIEM providers to evaluate their solutions in-depth.

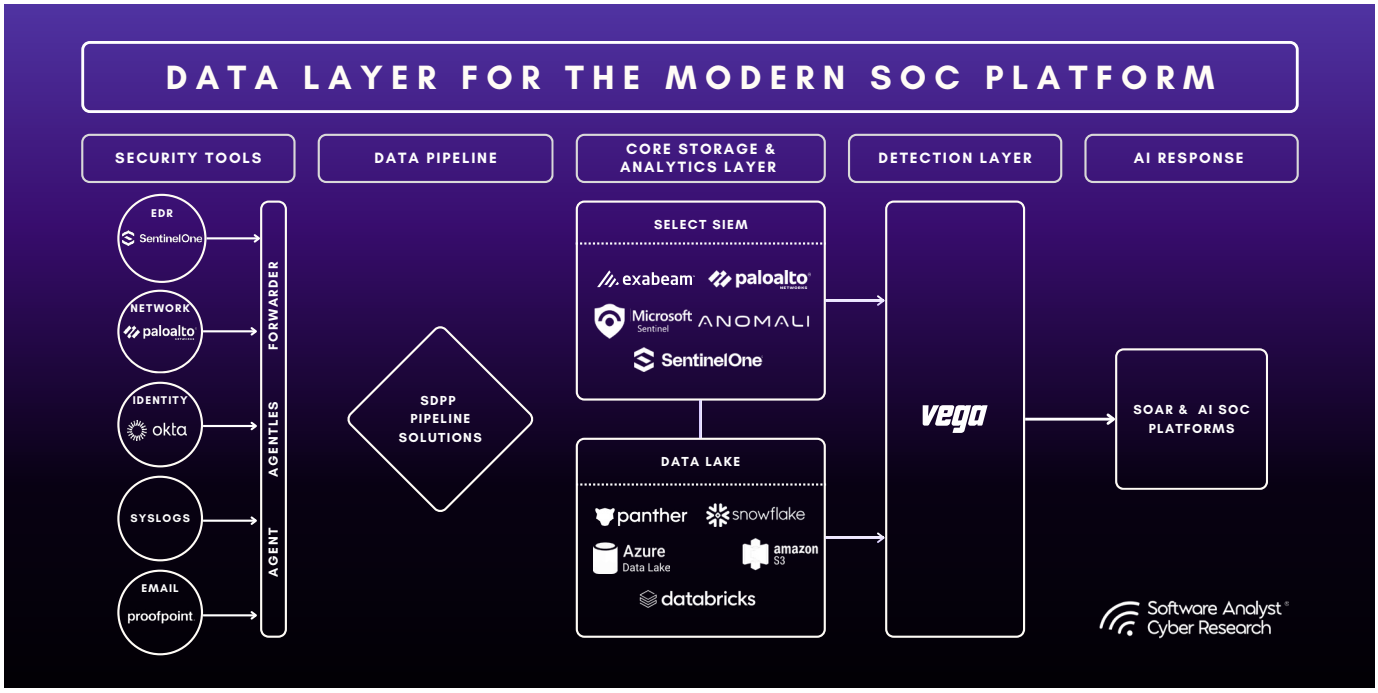
We conducted an in-depth analysis of many vendors. This report maps seven vendors to real-world use cases and SOC maturity, and uses our PDDIR framework (Pricing, Deployment, Detection, Investigation, Reporting) to compare how each tackles cost control, noise reduction, analyst efficiency, and openness. Backed by hands-on demos and questionnaires, it offers a practical decision playbook, clear selection criteria, and our opinion covering pipeline integration, pricing model shifts, AI for the analyst, and evolution of modern SIEM architecture.

Authors

Aqsa Taylor is the Chief Research Officer at Software Analyst Cyber Research, where she leads research initiatives and security leaders community.

Francis Odum is the Founder/CEO of the Software Analyst Cyber Research, where he leads the firm's research and engagement with cybersecurity leaders.

Rafal Kitab is a SOC and Incident Response leader at ConnectWise with extensive experience working as a Security Analyst, Engineer, Architect, Incident Responder and recently a Director. He brings considerable experience in Security Operations and shares his first hand experiences.



Introduction

Our Market Guide 2025: The Rise of Security Data Pipelines & How SIEMs Must Evolve and SACR AI SOC Market Landscape For 2025 reports set the stage for the major shifts we see in the Security Operations world, with SIEM at the center of it. Practitioners have been vocal about the rising costs, the operational overhead of rule management, and the alert fatigue with noisy logs. SIEM budgets are increasingly challenged by contenders such as **XDRs, security data lakes, security data pipelines, and other security analytics and operations platforms.** Yet one thing is certain: SIEM platforms are here to stand their ground. However, what defines a modern SIEM platform is now far more demanding.

What once began as a straightforward logging and analytics tool has transformed into one of the most complex and expensive platforms in the SOC. Vendors are advancing SIEM with deeper ties to data pipelines, AI-driven capabilities, modular designs, and a sharper focus on the analyst experience. What stands out in 2025 is how each leading provider is pushing the SOC forward in their own way.

This report highlights the major trends shaping SIEM in 2025 and what they mean for security teams. We look at how SIEM delivers value today, where it must evolve, and the characteristics that set modern platforms apart. The analysis focuses on how vendors are addressing long-standing pain points, how AI is changing the analyst experience, and how new architectures are reshaping the role of SIEM in the SOC.

To ground these themes, we assessed a set of vendors through demos and questionnaires using our comprehensive PDDIR (Pricing, Deployment, Detection, Investigation, Reporting) framework to show how these shifts are playing out in practice and what factors security leaders should consider when making decisions.

Actionable Summary

- SIEM is not disappearing, but transforming: Despite concerns with pricing and volume bloat, along with pressure from XDRs, data lakes, and security data pipeline platforms, SIEM remains central to the SOC; its definition has simply expanded.

- Cost and complexity remain top concerns: Rising data volumes and ingestion-based pricing models push buyers toward platforms that offer predictable costs, flexible storage, and reduced management overhead.

- Security data pipelines are reshaping SIEM: Filtering, in-stream detections, broader integrations, normalization and better data quality via SDPPs are now essential capabilities, helping to cut costs and reduce noise.

- Architectural shifts are accelerating: Vendors are moving from monolithic stacks to decoupled or distributed designs that separate compute and storage, enable federated queries, and give customers more control.

- AI is moving from hype to utility: Copilots, natural language detections, and automated investigations are being built into SIEMs, lowering the barrier for detection engineering and reducing analyst fatigue.

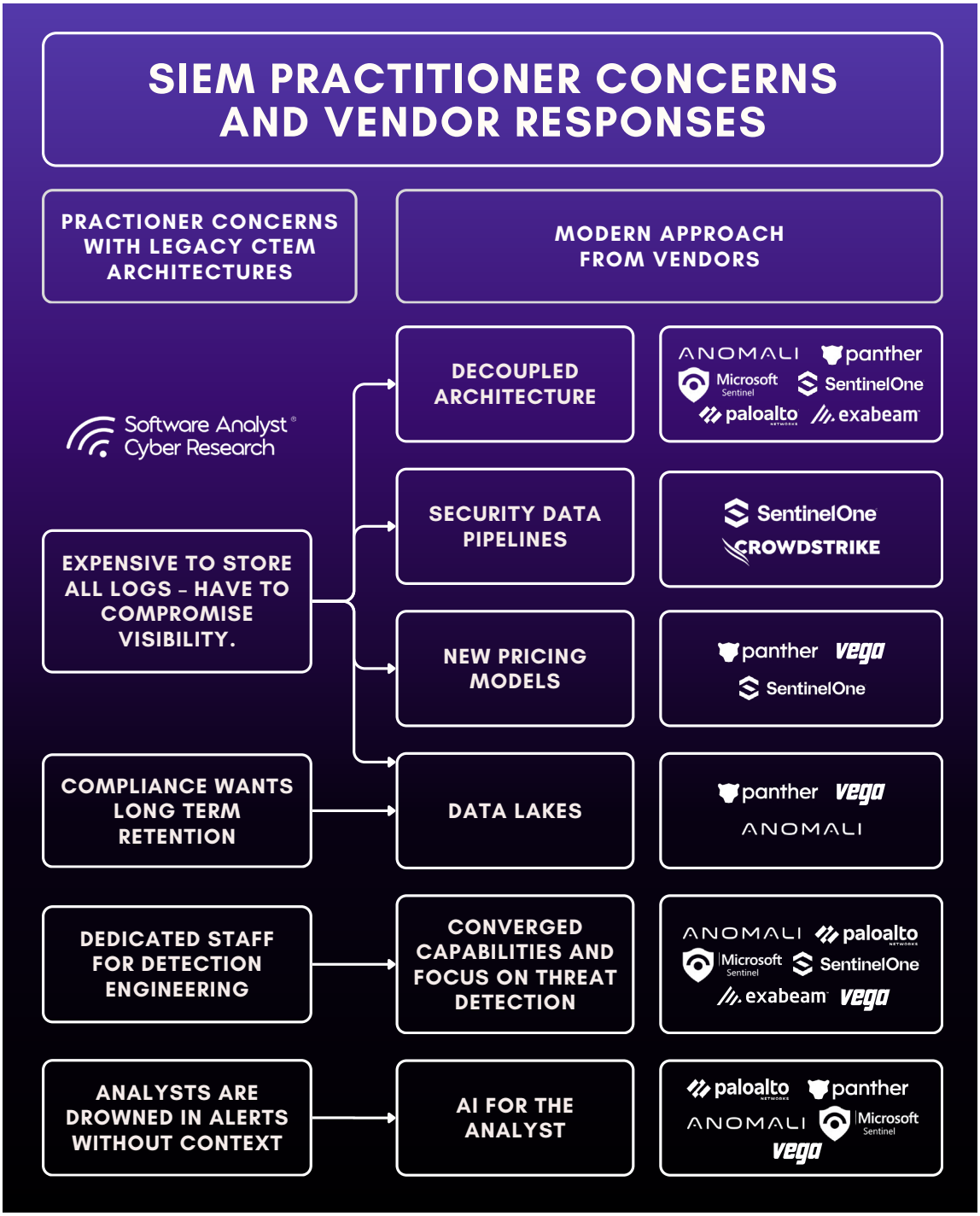
- Pricing innovation is critical: Many modern SIEM platforms are experimenting with pricing models based on data sources, utilization, or filtered events, in contrast to legacy ingest-based approaches.

- Convergence is reshaping the SOC: SOAR, XDR, exposure management and now SDPPs are being pulled into SIEM, making it less of a standalone product and more of a unified operating layer.

- Market split is emerging: SIEMs are evolving in two directions: open, decoupled overlays that maximize flexibility, and tightly bundled ecosystems that maximize integration.

From Legacy to Modern SIEM: Setting the Stage

Practitioners have made it clear: costs keep rising, data is noisy, and analysts are drowning in management overhead. Yet SIEM isn't going away. It's adapting! Taking on these challenges head-on with new architectures and smarter workflows. Although, in a modern SOC architecture, SIEM is no longer the stand alone player. It's becoming part of a bigger ecosystem, complemented by pipelines, data lakes, automation workflows, and adjacent platforms such as SOAR and XDR. To briefly summarize what led to the shift in modernization of SIEM platforms:



The Cost Problem

Data volumes keep rising as cloud adoption grows, and legacy SIEM pricing forces teams to choose between visibility and cost.

What organizations want:

- Noise reduction capabilities - help separate signal from noise
- Predictable pricing that avoids hidden cost bloats
- Full visibility without blind spots
- Long-term retention without penalties

The Overhead Trap

Legacy SIEMs are resource-heavy, demanding constant upkeep and specialized staff just to stay operational.

What organizations want:

- Faster onboarding with minimal manual setup
- Out-Of-The-Box content that reduces Time to Value
- Less reliance on dedicated detection engineering staff
- Automation that reduces day-to-day overhead
- Analyst time focused on investigation and response

The Analyst Dilemma

An industry research reveals that anywhere from 45% to 80% of security alerts are false positives. That is, 2 to 4 false alerts for every legitimate threat they uncover. It also states that analysts spend nearly three hours manually triaging them.

What organizations want:

- Reduced alert fatigue - better signal on threat
- OOTB threat detection content that keeps up with changing threat landscape
- Easy way to build custom threat detections
- Guided investigations with clear context
- Real reduction in repetitive, manual work

Market Shifts and Competitive Trends

In 2025, vendors are re-defining what a SIEM platform should look like by directly tackling the problems organizations have long faced with legacy tools. Some ways they are doing so:

Stronger focus on Security Data Pipelines

We reported earlier how Security Data Pipeline Platforms are quickly becoming a core component of the SOC architecture in our Market Guide 2025: The Rise of Security Data Pipelines & How SIEMs Must Evolve report. This has been proved true with the recent CrowdStrike's acquisition of Onum and SentinelOne's acquisition of Observo AI which highlight how vendors are moving toward this pipeline-first approach to address "noisy data" and cost concerns.

Coupling Security Data Pipeline Platforms (SDPPs) into the SIEM fabric elevates the platform beyond the traditional SIEM mold by addressing key shortcomings of legacy architectures -

1. **Filtering at ingestion –** Traditional SIEM platforms lack data quality controls at ingestion, leading to increased storage and analytics costs. Integration with SDPPs bridges this gap by filtering unwanted data at the source before it is saved in SIEM storage.
2. **Broader ingestion coverage –** SDPPs integrate with a wider range of data sources, helping expand the integration coverage of SIEM platforms.
3. **Detections in the pipeline –** By enabling in-stream detections, SDPPs significantly reduce Mean Time to Detect (MTTD) by avoiding storage indexes and processing delays.
4. **Cheaper storage options –** Some SDPPs include built-in data lakes and cold storage capabilities, providing more cost-efficient options for long-term data retention.
5. **Avoiding vendor lock-in –** SDPPs are built on open standards, allowing data to be routed to any destination. When paired with SIEM platforms, they can also simplify migrations.

In October, we will publish a report that takes a deeper dive into the SDPP market trends and insights.

Emergence of Decoupled SIEM Architectures

Historically, SIEMs provide a monolithic architecture without tiering - handling ingestion, normalization, and analysis in a single stack. Now, vendors are rolling out decoupled architectures that separate storage from compute, giving more flexibility and ownership of costs to the customers. Verdict is still out if such a move reduces pricing or makes it convoluted, however one thing is certain - it gives customers the flexibility to choose which was not an option before.

Data lakes are gaining popularity: Microsoft Sentinel Data Lake, SentinelOne's Singularity Data Lake, AWS Security Lake, and now Splunk's Machine Data Lake, to name a few. The goal is to address practitioners' pricing concerns by providing cost-effective ways to store data long term while still enabling analytics at reasonable speeds. At SACR, we view the market as being in a transitional stage: while data lakes are becoming more common, they are still most often coupled to a single analytics engine.

With Data lakes, come federated searches. We also see an emergence of the "query layer" model in reshaping how security teams think about data management in SIEM. Instead of forcing all telemetry into expensive analytics storage, some emerging platforms, such as Vega Security, are moving beyond this by building technology-agnostic SIEMs that can operate across data sources, no matter where the data resides. In an

ideal scenario, this creates an ecosystem where organizations can select the best-of-breed analytics platform alongside the best-of-breed data management solution. At minimum, it offers flexibility, allowing security data to be retained in any environment, including isolated or highly regulated ecosystems.

AI for the Analyst

AI is shifting from hype to utility. AI is being built into SIEM platforms at scale. The trajectory looks similar to Security Orchestration, Automation, and Response (SOAR), which started as standalone products before being absorbed into SIEM. Several vendors already demonstrated mature AI features embedded into their platforms.

Natural language processing is now a basic feature. More advance AI features include -

Agentic triage and investigations: AI summarizes alerts, stitches timelines, assigns risk, and suggests next steps. Improves MTTD and MTTR.

Guided builders and copilots: Step-by-step assistants for rule creation, hunt design, and case work. Speeds onboarding and reduces reliance on specialists.

Playbook synthesis and no-code response: Turn findings into executable workflows, with auto-generated or recommended actions. Accelerates containment and recovery.

This raises questions about the future: will dedicated AI SOC vendors remain a separate category, or will they be folded into SIEM like SOAR was? With natural language detection builders, copilots, and guided workflows, aimed at helping analysts create detections faster, investigate more efficiently, and spend less time on repetitive triage, AI is quickly becoming an expected integration in platforms. Check out our article on SACR AI SOC Market Landscape For 2025 to dive deeper into this conversation

New Pricing Models and Converged Capabilities

Other SIEM platforms are addressing the visibility vs cost problem by moving away from legacy ingest based pricing. We see models that are Integrations based, pay for filtered data, or custom in-house data lake models aimed at giving security leaders more predictable costs without forcing compromises on visibility or retention.

Capabilities like SOAR and XDR are have been pulled into the SIEM ecosystem for a long time and now we see even more cross convergence with tangential platforms, bringing in capabilities like UEBA and Data Pipelines. This convergence reduces tool sprawl and allows teams to detect, investigate, and respond within a unified workflow. Sentinel's deeper integrations with Microsoft Defender, Cortex XSIAM's integration with XDR, and similar moves from other vendors illustrate how the SIEM is evolving into a central operating layer for security operations, less a standalone product and more a platform where detection, investigation, and response converge.

Assessment methodology

For security leaders, the question is no longer whether to move past legacy SIEMs, but which modern approach fits their future. This report analyzes how vendors are redefining SIEM and provides insights to help CISOs and security teams navigate the shift and make informed decisions for the decade ahead.

We took a closer look at a handful of representative vendors, to anchor these themes in practical, real-world approaches. The vendors that we looked at (in alphabetical order) are:

- Anomali
- Exabeam
- Microsoft Sentinel
- Palo Alto Networks Cortex XSIAM
- Panther Labs
- SentinelOne
- Vega

To evaluate how these SIEM vendors are addressing the concerns from legacy SIEM approaches, we built a structured assessment process grounded in real operational needs and use cases. Each platform was measured against a broad set of criteria covering the most important functional, technical, and operational aspects of SIEM.

CATEGORY	FOCUS AREAS	KEY QUESTIONS/CONSIDERATIONS
Pricing	Addressing the legacy data quality and cost problems	How are they addressing the cost concern from practitioners?
Deployment	Flexibility and scalability	How are they deployed, how does their architecture benefit organizations?
Detection	Threat detection capabilities	How do they make detection easier? Do they use AI, and how much out-of-the-box content is included to keep pace with new threats?
Investigation	Analyst support	What built-in support is offered for analysts? How do they really improve the analyst experience? pre-written queries, natural language builders, or non-query approach?
Reporting	Operational reporting and monitoring	How does the platform provide meaningful risk assessment reports taking into account, threat context and business impact?

Software Analyst Cyber Research

Disclaimer:

The outcome of this assessment is not to place SIEM platforms in a tiered ranking, but to highlight how each is applying innovative approaches to address practitioner challenges and advance modern security operations.


ANOMALI

Anomali SIEM

Anomali SIEM, a core component of the Anomali AI-Powered Security and IT Platform, differentiates itself from traditional platforms by placing threat intelligence at the core of its design. Instead of focusing only on log collection and correlation, it enriches data with threat intelligence context, helping analysts prioritize detections tied to relevant threat actors, industries, or geographical regions. This functionality is delivered via the Open Data Lake, which is included with the platform. They also offer a decoupled architecture - where compute and storage modules are separated.

Another differentiator is their speed in retrieving data, even if stored long-term (7+ years). Practitioners have often raised concerns about the time required to retrieve data from legacy platforms, for historical investigations. With Anomali, there is no tiering in their storage; all data is stored as hot (seven+ years). During the demo, we saw it retrieve years of data within seconds.

On the detection side, Anomali introduces the Anomali Query Language (AQL), supported by Anomali AI that translates natural language into structured queries, aimed at lowering the learning curve for analysts. In addition to the NLP function, Anomali AI can also extract IOCs from webpages, assist in predictive analytics, and summarize complex logs and intel briefs for rapid analysis & consumption. Content can be tailored during onboarding to match industry-specific threat patterns, and MSSPs can manage multiple tenants through the Anomali Platform. Anomali also reports very fast search performance across billions of records, allowing rapid pivoting and investigation at scale. Taken together, these capabilities position Anomali as a SIEM option focused on threat intelligence, cost efficiency, and speed.

Cost Effectiveness

Anomalis' decoupled storage and compute architecture avoids the traditional ingestion-based pricing traps that often force clients to choose which data is "worth" keeping. Instead, all data can be ingested into the open data lake, including non-security sources, without immediate financial pressure. Pricing is positioned as 40–60% lower than comparable SIEMs, and because all data remains "hot" for seven+ years, customers avoid surprise retention costs. This approach allows organizations to focus on visibility.

Deployment

The platform supports flexible deployment across SaaS, hybrid, air gapped, and on-premises environments, with government customers often favoring the latter. Out of the box, Anomali integrates with over 125 standard log formats and most mainstream firewalls and endpoint security solutions, while also offering custom parsers where needed. Data collection is enriched by Anomali's threat intelligence "Match" function, which automatically correlates log data with known malicious indicators. Integrations extend downstream as well, enabling automated responses such as pushing high-confidence IoCs directly into EDR or firewall systems for blocking. These are prioritized through a composite confidence score, which blends Anomali's own intelligence with feeds from partners like CrowdStrike, Recorded Future and others. The platform is based on an open, unified data lake. Data is stored in open formats, which limits vendor dependency and allows scaling flexibility. The data lake remains under the customer's ownership and governance, including storage and access policies.

Detection

During onboarding, customers receive tailored content (dashboards, alert rules, and pre-built queries) aligned with their specific industry and region. Threat actor profiles and campaign intelligence are integrated into detection engineering, helping teams to focus on the most likely attack patterns. Analysts interact with detections primarily through AnomaliQuery Language (AQL), which is designed for scale and

speed, but can be accessed more intuitively via the Anomali AI that translates natural language questions into AQL queries. Queries can be saved, reused, and converted into alerts with flexible look-back periods and throttling options. Although detection-as-code is not yet available, JSON-based content deployment is supported today, and MSSPs gain additional leverage through the platform for managing detections across multiple tenants.

Investigation

Investigation workflows are enhanced by AI and contextual intelligence. The Anomali AI guides analysts by generating or refining queries, summarizing alerts, and accelerating triage. The system enriches alerts with MITRE mapping, event entity details, and known threat actor context, giving analysts a head start on investigations. Collaboration is supported through features like "trusted circles," which allow organizations in parent-child or partner relationships to share indicators and reports. Together, these features reduce time spent on manual enrichment and enable analysts to work from a higher-value starting point.

Reporting

On the reporting side, Anomali's architecture allows for both granular searches and broad historical queries, with nearly infinite horizontal scaling. This ensures SOC teams can track metrics over time without being restricted by storage limits or forced into archival tiers. Operational health is managed through efficient storage mechanisms that minimize data usage while keeping all data immediately accessible. The result is a platform that supports both the compliance-driven need for long-term retention and the operational reality of quick look-backs during investigations.

Strengths

- Strong focus on threat intelligence as the foundation of the platform, enabling more relevant detections.
- Probably the fastest search we've seen
- Decoupled storage and compute architecture eliminates ingestion-based pricing constraints.
- Flexible deployment options across SaaS, hybrid, Air-Gapped, and on-premises.
- Tailored onboarding content aligned to industry and geography.
- Open Unified Data Lake, supporting open formats

Areas to Watch

- Detection-as-code is not yet fully supported (currently only JSON-based deployment).
- Market presence is smaller compared to Microsoft and Splunk, meaning fewer community-created detections and integrations.
- Although storing all data in hot storage can improve retrieval speeds, it raises cost concerns as hot storage can be more expensive, compared to cold storage. Anomali's approach to addressing this concern is with their open data lake model and decoupled architecture.

Conclusion

SIEM Market Evolution – SACR Prediction

Looking ahead, our team predicts that the SIEM market is likely to evolve along two distinct paths -

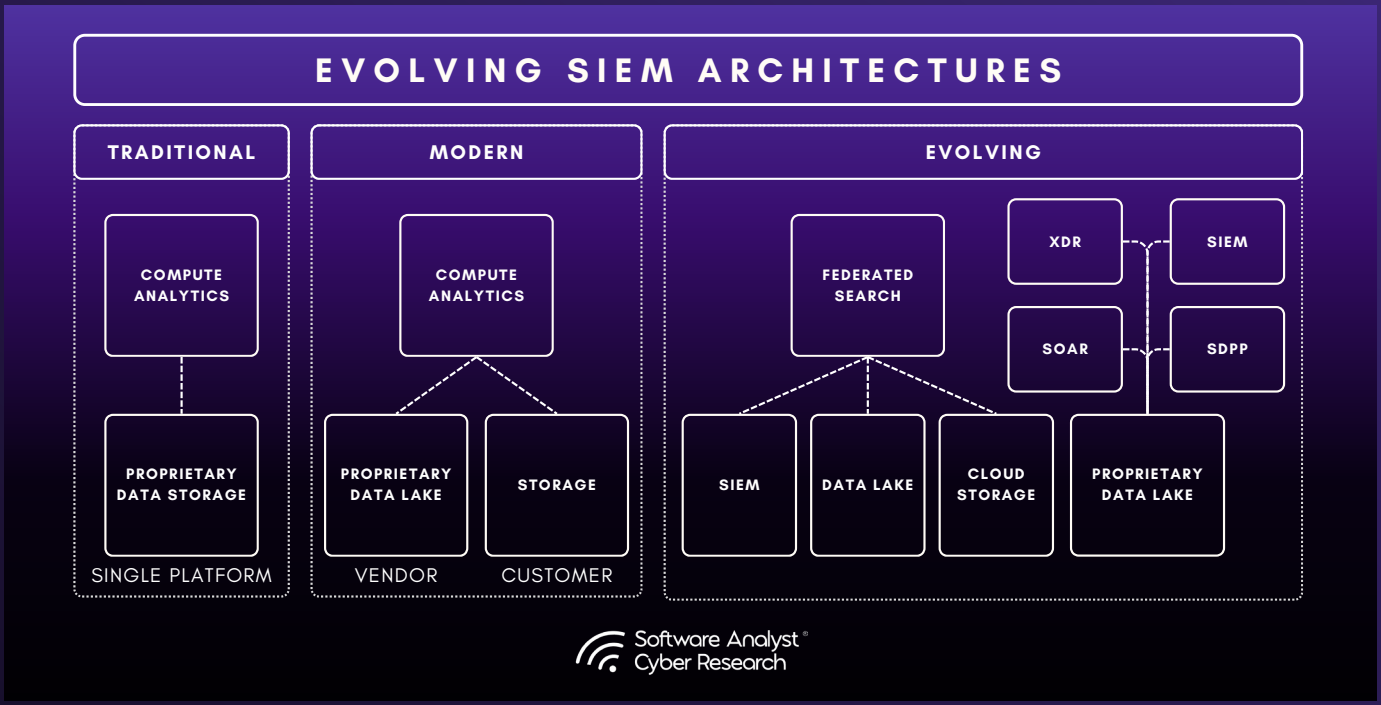
Decoupled Distributed Architectures

The first path centers on the emergence of decoupled, technology-agnostic SIEMs. These platforms will operate as a security analytics overlay that can query and analyze data wherever it resides, whether in a vendor's data lake, object storage, or third-party environments. This approach provides flexibility, cost control, and choice. Some vendors are already pursuing this model, creating federated detection layers that decouple analytics from storage and reduce dependency on a single stack. Over time, such architectures may enable organizations to combine best-in-class analytics with best-in-class data management, unlocking both efficiency and innovation.

Convergence and Bundling with SDPP

The second path is one of convergence and bundling. Here, SIEMs increasingly become the central pillar of a broader security ecosystem, tightly integrated with security data pipelines platforms (SDPP), endpoint, identity, cloud, and network security tools. This strategy is exemplified by moves from CrowdStrike, Palo Alto Networks, and Microsoft, where SIEM capabilities are packaged with XDR, SOAR, and adjacent security offerings. For customers, this creates a coherent ecosystem with seamless integration, unified workflows, and simplified vendor management, at the cost of flexibility and greater dependence on a single provider.

Both directions respond to practitioner pain points but offer contrasting value propositions: openness and choice versus consolidation and cohesion. The market will likely continue to fragment along these lines, with buyers aligning based on their organizational maturity, risk appetite, and technology philosophy.



The future of SIEM is not about whether the technology survives. Rather it is about **how it reinvents itself.** Whether through decoupled architectures that provide flexibility in choosing where the compute vs storage act as analytics overlays across distributed data, or through tightly bundled ecosystems that unify the entire security stack, SIEM will remain central to the SOC. The question for security leaders is which path aligns with their vision: maximum flexibility and independence, or maximum integration and simplicity.

Software Analyst®
Cyber Research

business

personal

softwareanalyst.io

SACR®

Trusted research. Sharp insights. Real conversation.

CISO	VENDOR
SECURITY TEAMS	INVESTORS