The Decoupled SIEM:

An Architectural Foundation for Agentic SecOps

AUTHOR

Oliver Rochford

Cyber Futurists

August 2026

Executive Summary

The SOC's scaling problem is no longer primarily a shortage of storage or tools. Core workflows like onboarding data, searching across sources, developing detections and investigating alerts still depend heavily on manual work. Telemetry volumes and environmental complexity are growing faster than many teams can absorb. Adding another console or changing SIEM vendor does not fundamentally change that operating model.

Agentic SecOps meets these challenges by changing the division of labour. AI agents execute bounded portions of recurring workflows, while people set policy, approve consequential actions and handle ambiguous cases. While an assistant recommends work, an agent performs defined work within explicit permissions, using evidence and producing an auditable record of its actions.

But agents also introduce a practical access problem. Agents need governed access to telemetry and context distributed across SIEMs, data lakes, cloud platforms, identity systems and endpoint tools. A decoupled SIEM separates the pipeline, storage, analytics, response and agent layers so that each can evolve independently. Integrated SIEM platforms can expose similar capabilities through data lakes, APIs and federated queries. Vendor count matters less than whether teams can reach the evidence they need, move their detection content and constrain automated actions.

Decoupling is a spectrum rather than a rip-and-replace programme. Organisations can extend an existing SIEM with a lake for selected telemetry, use a shared data layer to bridge separate security ecosystems, or build a modular stack without a legacy SIEM. Each approach can deliver value before the organisation introduces agentic workflows. Published Anvilogic customer cases report savings of 70–90% in selected deployments, but these do not establish a general TCO benchmark: compute, migration, engineering and dual-running costs still need to be modelled.

For CISOs, the starting point is the work the organisation intends to automate and whether the current architecture can support it. Can agents reach the required evidence? Can detections and workflow definitions move between platforms? Are actions constrained, auditable and reversible? Can one component be changed without rebuilding adjacent layers? If those conditions are absent, adding an agent will not correct the underlying data and workflow problems. If they are present, the organisation can automate selected work while retaining control and accountability.

KEY FINDINGS

1 Manual security operations no longer scale

Telemetry volume, alert counts and environmental complexity are growing faster than many organisations can add specialist capacity or capability. Detection content is difficult to author and maintain at the pace infrastructure changes, and analyst time is often consumed by triage that produces little durable improvement. Hiring can relieve individual bottlenecks but does not remove the fundamental mismatch between growing workload and finite human attention.

2 Agentic SecOps is the emerging operating model

Assistants that summarise the alert queue address only part of the workload. The next stage is AI agents that execute bounded SOC workflows: onboarding data sources, running federated searches, building and tuning detections, and triaging and investigating alerts. Human approval remains necessary where judgment or consequential action is involved, and the system should preserve an auditable record of evidence, tool calls, actions and outcomes. Gartner placed AI SOC agents at the Innovation Trigger stage with reported market penetration of 1–5% in its 2025 Hype Cycle for Security Operations. A sensible first step is to test one bounded workflow and measure its accuracy, operating cost and demand for human intervention.

3 AI readiness depends on data architecture

MIT NANDA reported that 95% of the enterprise GenAI deployments it examined produced no measurable P&L impact. Its principal explanation was a learning and workflow-integration gap rather than data readiness alone. In the SOC, agents additionally require complete, normalised telemetry and contextual enrichment from asset, identity and threat-intelligence sources. Open, API-driven access reduces one important constraint, but reliable operation also depends on evaluation, evidence, permissions and governed workflows. Better data removes one failure mode; it does not supply evaluation, permissions or workflow design.

4 The cost crisis has shifted from ingestion to compute

Traditional SIEM economics have often been shaped by ingestion volume. Data lakes can reduce storage costs, but AI-driven detection, cross-silo correlation and agentic investigation create new compute demand on the analytics layer. The cost question therefore extends beyond storing telemetry to the frequency and complexity of analysis performed against it. Organisations that separate storage from analytics still need to model query compute, egress, engineering and operational overhead. Vendor acquisitions and packaging changes may add commercial uncertainty, but the architectural case should stand independently of any single transaction.

5 Decoupling works as a spectrum

Many organisations do not replace their SIEM; they extend it. Three entry patterns recur in vendor and customer examples. In the Extend pattern, a data lake is added alongside the existing SIEM for telemetry that is uneconomic to ingest there. In the Bridge pattern, the lake becomes a shared correlation layer across otherwise disconnected security ecosystems. In the Build pattern, MSSPs, MDRs and cloud-native organisations design composable architectures without a legacy SIEM constraint. These are entry patterns rather than a mandatory sequence, and each creates different requirements for data engineering, detection content and governance.

6 Detection engineering: from craft bottleneck to automated workflow

A general-purpose data lake does not provide the parsers, detection rules, correlation logic or response playbooks expected from a SIEM. Historically, the scarcity of teams able to operationalise Detection-as-Code was a major constraint on decoupling, and capability investment still matters. Detection engineering is now becoming one workflow within a broader automated lifecycle: onboard, search, detect and investigate. Agentic platforms can reduce the effort required to draft, translate and test detections, although deployment speed depends on schema quality, validation and approval. Organisations that treat the lake as cheaper storage without investing in detection content and ownership will still struggle to produce operational value.

1. The Problem: Security Operations is hitting human limits

Ask a SOC manager what has changed in the past five years and the answer is a list of "mores": more data, as cloud workloads, SaaS applications, containers, EDR telemetry, and identity systems multiply log volume; more alerts, as detection surfaces expand faster than tuning capacity; more distributed environments, as security data spreads across SIEMs, data lakes, object stores, and a dozen cloud and identity platforms; and more investigations, each demanding context scattered across those same silos. What has not changed is the team. Security headcount has grown incrementally where it has grown at all, and the skills required (detection engineering, data engineering, cloud forensics) are among the scarcest in the industry.

The result is a structural mismatch. Every workflow in the traditional SOC is manual at its core: humans onboard data sources, humans write and tune detection rules, humans triage alerts, humans assemble investigation context from five consoles. Automation to date has meant SOAR playbooks bolted onto the edges of this model, scripted responses to known scenarios, while the substance of the work remains human-paced. The workload is growing at machine speed and the operating model still runs at human speed.

The Legacy: A Platform Built for a Different Era

When SIEM emerged in the early 2000s, the design brief was simple: centralise data from multiple intrusion detection and prevention systems, correlate events, and surface alerts. The log sources were manageable, the data volumes were modest, and the value proposition was clear.

Two decades on, the SIEM has expanded far beyond its original scope. Log management, compliance reporting, UEBA, SOAR, threat intelligence and AI-assisted triage have all been added or absorbed. Some current SIEM platforms also incorporate data lakes and federated access to external stores, so the distinction is no longer simply between one product and several products. The relevant question is whether storage, analytics, detection content and response can evolve independently, and whether the organisation can access and move its data and logic without prohibitive switching costs.

The Budget Wall

The economics compound the scaling problem. SIEM vendors charge per gigabyte ingested, and ingestion has shifted from a customer success metric to a revenue driver. Cloud-scale telemetry volumes have grown exponentially while SIEM budgets have not kept pace.

Organisations are forced to choose which telemetry to ingest and which to discard, based not on security value but on cost constraints. Full EDR telemetry — the raw process execution data that enables deep forensic investigation — gets discarded because the per-gigabyte cost of ingesting it into a SIEM is simply too high.

Historically, comprehensive ingestion into a high-cost analytics tier was practical only for organisations with very large security budgets. Even then, additional data did not guarantee effective detection or response if analytics, content and operating capacity failed to keep pace.

Industry consolidation, including Cisco's acquisition of Splunk, has prompted some customers to reassess pricing, packaging and platform-direction risk. These concerns may accelerate architectural reviews, but they are commercial triggers rather than evidence for decoupling on their own.

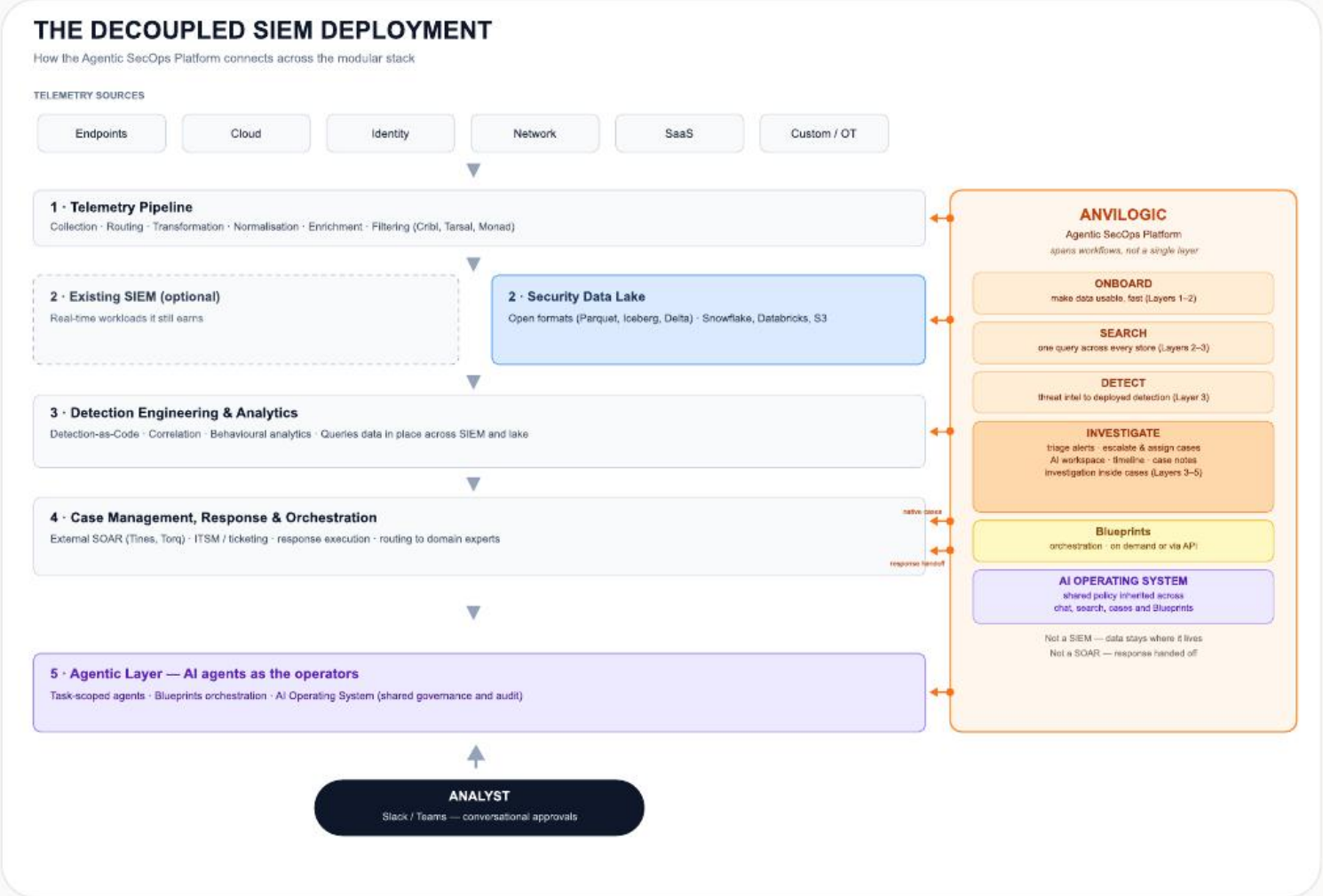
The Next Wall Is Compute — and Labour

Ingestion cost is no longer the primary constraint. Organisations that moved telemetry to a data lake have addressed storage economics. But AI reshapes the cost model: running behavioural analytics, cross-silo correlation, and agentic investigation against a petabyte-scale lake is compute-intensive and significantly more expensive than rule-based detection. The cost question has shifted from left to right, from "can I afford to ingest this data" to "can I afford to reason over it?"

Lower-cost storage makes broader retention possible, but it does not reduce the human effort required to use that telemetry. Organisations need to change both the architecture and the way work is assigned between tools, agents and people.

2. What Is Decoupled SIEM?

Decoupled SIEM is an architecture pattern, not a turnkey product. The traditional SIEM stack is decomposed into independent, interchangeable layers, each responsible for a distinct function and connected by open pipelines rather than proprietary integrations.



In a tightly integrated SIEM, collection, normalisation, storage, detection and response are managed within one commercial platform. This can reduce integration effort, but it may also couple migration of historical data, schemas, enrichment logic and detection content. The resulting switching cost depends on export options, APIs, contractual terms and the portability of the organisation's own content.

In a decoupled SIEM, each function becomes a semi-independent component. The organisation, not the vendor, controls the pipeline. Data is stored in open formats in a security data lake, and the analytics layer operates independently of storage, querying data in place rather than requiring ingestion into a proprietary store.

Dimension	Monolithic SIEM	Decoupled SIEM
Architecture	Collection + Normalisation + Storage + Detection + Response in one platform	Each function operates as a separate, modular component
Pipeline control	Vendor controls the pipeline	The organisation controls the pipeline
Data format	Data commonly stored in vendor-managed formats	Data can be stored in customer-controlled or open table formats when deliberately designed that way
Analytics coupling	Analytics tied to storage and proprietary data formats	Analytics layer operates independently of storage
Detection ownership	Rules commonly tied to vendor schemas and query languages	Rules owned by accountable authors and version-controlled; portability still requires translation and validation
Alert routing	Centralised SOC queue; tiered L1/ L2/L3 escalation	Decentralised routing to system owners and domain experts
Response model	Manual triage and escalation chains; SOAR added on	Automated playbooks for known scenarios; human escalation as needed
AI role	Vendor-locked analytics; AI constrained to one data model	Composable agentic layer with open access across the full stack
Vendor dependency	Potentially high: migration may require data, schema and content translation	Lower when interfaces, data and detection content are genuinely portable

Five Decomposed Layers

The architecture is organised into five functional layers. The first four are technical infrastructure; the fifth sits at the interface between the human and the technology stack, where AI agents mediate the interaction.

1. Telemetry Pipeline. In a monolithic SIEM, collection and transformation are tightly coupled, with the vendor controlling ingestion, parsing, and schema. Modern telemetry pipelines decouple these functions. Tools like Cribl, Tarsal, and Monad handle collection, routing, transformation, normalisation, enrichment, and filtering in a unified pipeline layer. They are source-agnostic for ingestion and destination- and use-case-agnostic for output. They solve the problem of getting data from A to B and making it usable.

2. Storage. Security data lakes and data warehouses such as Snowflake, Databricks, or cloud-native object stores replace the proprietary data store at the core of the traditional SIEM. Data lands in open formats (Parquet, Iceberg, Delta Lake) and is queryable by any tool with the appropriate connector. That openness is the architectural linchpin: multiple layers can query the same lake simultaneously. Detection engines, AI agents, threat hunters, and compliance teams all operate on the same data without duplicating it into separate silos.

Storage in a data lake can be materially less expensive than equivalent SIEM ingestion and retention, although the difference depends on storage tier, query frequency, egress and engineering overhead. This allows organisations to retain substantially more telemetry and choose different tiers for active detection, investigation, compliance and long-term retention. Storage becomes a broader design choice, but it does not cease to be a budget constraint.

3. Detection Engineering and Analytics. Detection logic in an integrated SIEM is commonly tied to the vendor's query language, data model and schema. Rules written in SPL, KQL or other platform-specific languages can be versioned and tested, but moving them between platforms usually requires translation, field mapping and renewed validation.

In the decoupled model, detection engineering becomes a first-class discipline. Correlation engines, detection rules, behavioural analytics and threat-detection logic can run as an independent layer above the data lake. Detection-as-Code allows rules to be managed in version control, tested through CI/CD and deployed programmatically. Portability improves at the level of intent, metadata and lifecycle, but executable logic still requires field mapping, query-language translation and platform-specific validation before it can run across Snowflake, Databricks or a traditional SIEM.

Two operational principles make this layer work. First, every detection rule has a named owner accountable for its performance — someone specifically responsible for false positives or missed patterns. Second, mature detection engineering pairs each rule with a defined response: not just "alert fires," but the associated triage steps, enrichment lookups, and escalation path. Detection without a response plan is noise.

4. Response and Orchestration. The operational model diverges most sharply from the traditional SOC here. In a monolithic SIEM environment, every alert lands in a centralised queue. Tier-1 analysts triage and escalate to Tier-2, which in turn escalates to Tier-3. The person who resolves the issue is often three handoffs away from the original alert.

A decoupled architecture can support a more distributed response model, but it does not require one. Alerts may be routed to system owners and domain experts who can act, while SOAR platforms and automated playbooks handle well-understood scenarios such as isolating an endpoint or disabling a leaked credential. Collaboration tools can provide a coordination interface, but case management, evidence retention and escalation policy remain necessary. Alert routing should reflect who may act and who remains accountable for the result.

5. The Agentic Layer: AI Agents as the Operators. This layer is new, and it does not fit neatly into the traditional SIEM stack because it is not a conventional technology layer. It sits between the analyst and the technical stack, mediating the interaction between human and machine — and increasingly doing the work itself.

AI agents in this layer can query the data lake, invoke detection logic, enrich findings with identity, asset and threat-intelligence context, correlate across sources and conduct bounded investigations. Results may be delivered through a case-management system, collaboration tool or analyst interface. The interface is secondary to the operating controls: scoped access, explicit approval points and a complete record of evidence and actions.

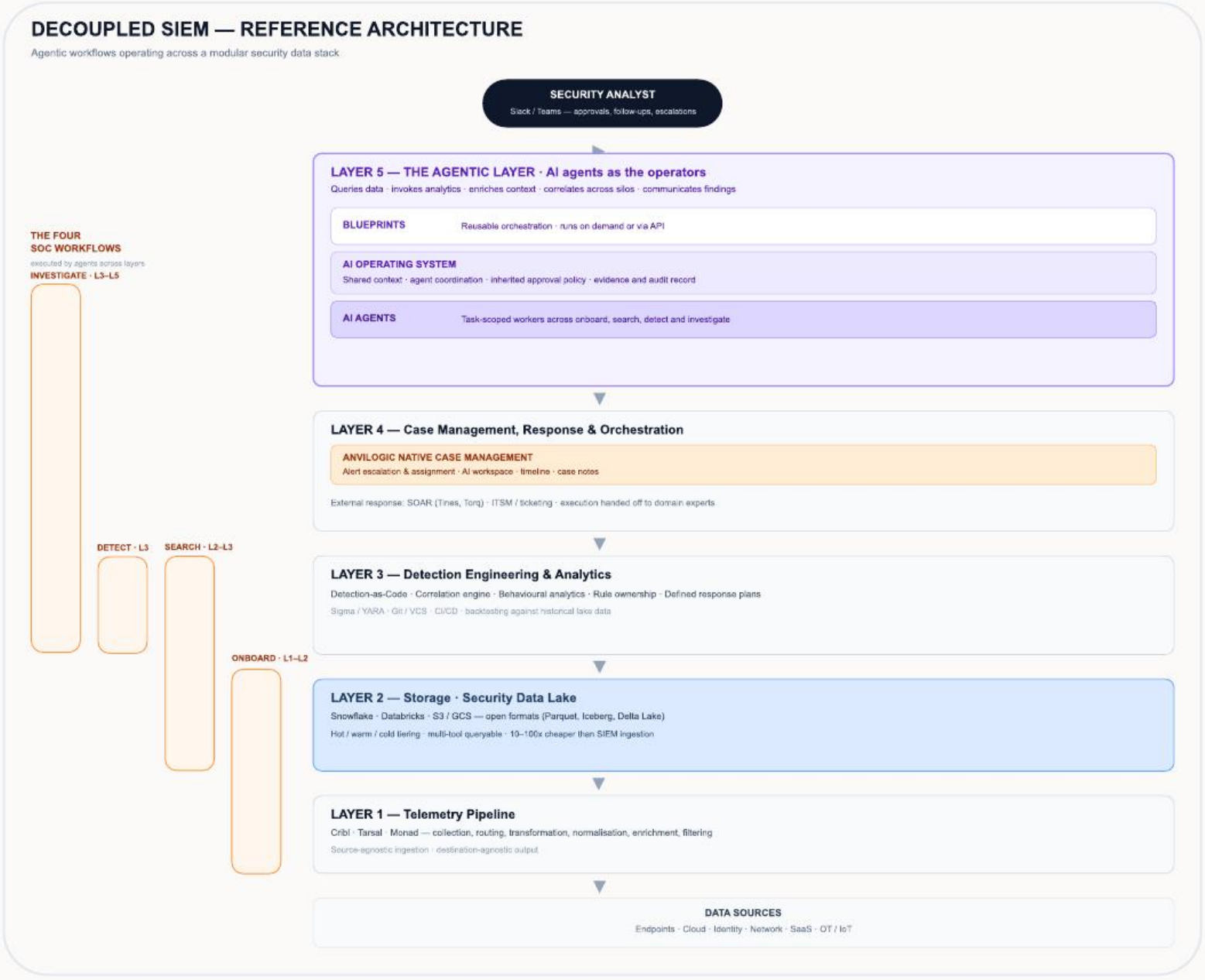
Two properties separate agentic operation from AI assistance. First, the agent performs bounded actions rather than only advising: for example, it can deploy a detection, run an investigation or escalate a case. Second, it operates under explicit governance. Approval gates remain where human judgment or consequential action is required, while the system records its inputs, retrieved evidence, tool calls, actions and outcomes. Reusable workflow definitions can preserve operational knowledge and make these controls repeatable; Anvilogic calls its implementation Blueprints.

Agents often need controlled access across several layers of the stack. An integrated SIEM can expose APIs, a data lake and federated sources, but the organisation remains dependent on the breadth, portability and commercial terms of those interfaces. Decoupling reduces that dependency when components and data can genuinely be exchanged.

Operating Model and SOC Workflows

A useful way to understand what runs on this architecture is by workflow rather than by layer. For the scope of this paper, four recurring analytics workflows show how work moves across the layers.

They do not represent the whole of security operations: response execution, case management, governance and assurance remain separate responsibilities.



Workflow	What it means	Where it runs
Onboard	Parse, normalise and map new data sources so they are usable — without a data engineering project	Layers 1–2 (pipeline and storage)
Search	Query everything from one place, across SIEM, lake and object storage, without moving the data	Layers 2–3 (federated across storage and analytics)
Detect	Turn threat intelligence into validated, deployed detections across whatever platforms hold the data	Layer 3 (detection engineering and analytics)
Investigate	Deliver triage, enrichment and severity assessment before a case is opened — context, not raw alerts	Layers 3–5 (analytics, orchestration and agents)

Anvilogic, the sponsor of this research, positions its platform around these four workflows across an existing stack that may include Splunk, Snowflake, Sentinel, Databricks and S3. It can operate on a data lake or alongside a SIEM, while handing response execution to SOAR and ticketing tools. This illustrates a cross-layer implementation rather than a complete replacement architecture. Other products draw the boundaries differently, and integrated SIEM platforms increasingly combine data-lake, federation, analytics and response capabilities in one commercial platform.

Origins and Lineage

The decoupled SIEM pattern was established by practitioners as well as vendors. Netflix's 2018 SOCless model was an influential early example: security operations built on independent AWS services, a data lake and automated analytics. The pattern has since appeared in other cloud-native environments seeking greater control over scale, cost and workflow design.

Terminology has also evolved. SOCless, SIEMless, Decoupled SIEM, and Deconstructed SIEM are used interchangeably. A related concept is the headless SOC, the security equivalent of a headless server. AI agents collect and analyse telemetry, deliver structured summaries with findings via Slack, and allow analysts to respond conversationally without interacting with a traditional dashboard. Agentic SecOps is the operating model that emerges when these threads converge: modular architecture underneath, governed agents on top.

Fortune 500 Financial Services Firm

An Anvilogic case study describes a Fortune 500 financial-services organisation ingesting 1 TB per day of CrowdStrike Falcon Data Replicator telemetry into Snowflake, reaching 7 TB of raw events during the trial. Anvilogic reports that the team deployed 75 use cases, saved $1.2 million compared with its modelled alternative and required one hour of Anvilogic setup time to begin deploying content. These are vendor-published case-study results rather than an independently verified benchmark.

3. Why Now? The Convergence of Three Pressures

Three pressures are making the manual, monolithic model untenable for a growing segment of the market while simultaneously enabling the agentic one.

3.1 Manual Work No Longer Scales

The platform consolidation debate deserves more nuance than the marketing suggests, but it is ultimately a debate about the wrong constraint. Many organisations pursue a platform strategy for valid reasons: fewer vendors, fewer contracts, less integration overhead. Yet consolidating tools does not consolidate work. The detection backlog, the alert queue, and the investigation load remain human-paced regardless of how many logos are on the invoice.

Cybersecurity is also horizontal, not vertical. A retailer with distributed POS endpoints needs XDR. A bank running custom finance applications needs a security data lake and insider-risk analytics. A manufacturer with OT environments needs network-centric detection. One size does not fit all, and stitched-together mega-suites assembled through M&A rarely provide the depth complex organisations require. What every one of these profiles shares, however, is the same bottleneck: the number of skilled humans available to onboard sources, write detections, and run investigations.

The meaningful shift, then, runs from manual operations to automated ones rather than from best-of-breed to platform. Organisations need to automate security operations as well as modernise infrastructure; modernisation on its own produces a cheaper place to do the same manual work. The modern SOC's interface is already Slack or Teams; analysts spend their working hours in communication tools, and adding another dashboard provides no value. The decoupled model aligns with this: analytics and agents operate in the background, while the interface exists wherever the analyst already is.

3.2 AI Readiness: AI Alone Is Not Enough

The security industry is investing heavily in AI agents for the SOC, and major SIEM and XDR vendors have introduced AI-assisted capabilities. Adoption remains early. Gartner's 2025 Hype Cycle for Security Operations placed AI SOC agents at the Innovation Trigger stage with reported market penetration of 1–5%. MIT NANDA separately reported that 95% of the enterprise GenAI deployments it examined produced no measurable P&L impact. The two findings describe different populations and should not be treated as a single adoption statistic.

MIT NANDA attributed the enterprise divide principally to a learning and workflow-integration gap. Data readiness is an additional constraint in the SOC: incomplete, inconsistent or poorly enriched telemetry limits the evidence available to an agent. Clean data improves the conditions for reliable operation, but it does not eliminate model error or guarantee accurate verdicts.

Agent performance depends on the evidence and controls around the model. Agents need governed access to the required telemetry, contextual enrichment from asset, identity and threat-intelligence sources, and defined workflows rather than an open-ended mandate. Decoupling can allow data and analytics to evolve independently, but it is not the only way to expose external data: current SIEM platforms increasingly support data lakes, APIs and federated queries. Reliable operation additionally requires evaluation, evidence, scoped permissions, action controls and rollback. Evaluate each option against the work it must support: reaching the necessary evidence, moving detection content and constraining automated actions.

3.3 Data Economics

Cloud-scale telemetry is growing faster than SIEM budgets can support. Consider the complete Falcon Data Replicator (FDR) feed from CrowdStrike: every process execution, network connection, and file operation across an endpoint fleet. Ingesting this at per-GB pricing is prohibitively expensive.

Data lakes can change the economics. Anvilogic customer case studies report savings of 70–90% in selected deployments compared with equivalent SIEM arrangements. Those cases demonstrate potential rather than a market average, and their calculations should be read alongside query-compute, migration, engineering and dual-running costs. Where savings are realised, they may fund further analytics and automation, including agentic initiatives.

Alteryx

Alteryx used Anvilogic as a bridge while transitioning from a legacy SIEM, separating detection content from the underlying storage decision. Alteryx's Sr. Director of Security Operations and Engineering said the platform's schema abstraction allowed the company to plan for future storage changes without committing immediately to one data lake or SIEM. The published case study does not provide the 70–80% cost-reduction figure sometimes associated with a separate Anvilogic customer example.

1 Gartner, Hype Cycle for Security Operations, 2025. 2 MIT NANDA, The GenAI Divide: State of AI in Business 2025.

4. Decoupled SIEM Is a Spectrum

Decoupling is neither all-or-nothing nor strictly linear. How organisations enter the spectrum matters as much as where they sit on it.

Three entry patterns dominate, and they correspond directly to the deployment stories a security leader will hear from vendors in this market. Extend — bolt a data lake on alongside the existing investment to capture telemetry the SIEM cannot economically ingest; in vendor language, augment what is already there. Bridge — use the data lake as a common substrate across disconnected security ecosystems, often shifting high-volume telemetry to the lake while the existing SIEM stays live on the workloads it earns; in vendor language, modernise. Build — design composable from day one, with no legacy to decouple from; in vendor language, run standalone. The entry pattern determines where you start. Operational maturity determines how far you go.

The Decoupling Spectrum

Level	Description	Example
0 — Monolithic	Traditional all-in-one SIEM	Splunk handling everything
1 — Data-Decoupled	Data lake for retention, SIEM for analytics	Data lake for long-term storage + SIEM for real-time analytics
2 — Analytics-Decoupled	Separate analytics layer on top of the data lake	Anvilogic running detection and correlation across SIEM and data lakes
3 — Fully Modular	Every layer is independently swappable, pipeline-first; agentic workflows are viable across the full substrate	Full SOCless / headless SOC model

5. Profile of the Successful Adopter

Not every organisation should adopt a decoupled model.

Profiles That Succeed

Cloud-native organisations operating under shared responsibility models. This is where the decoupled model originated. Teams running web, middleware, and database stacks on AWS or Azure, where development, operations, and security share incident response, have infrastructure-centric telemetry, and think in pipelines and APIs. Their response model is "fix the code, not apply a patch", requiring developers or system integrators, not Tier-1 analysts. Netflix's 2018 SOCless architecture[3] fits this profile: no traditional 24/7 SOC or monolithic SIEM, but pipeline-driven detection, automated response, and escalation to system owners.

Large enterprises with existing data lake investments. If Snowflake or Databricks is already in production for BI, product analytics, or data science, the incremental cost of extending an existing data lake to security telemetry is significantly lower than building from scratch. The security team inherits both the platform and its capabilities.

Regulated financial institutions. Banks face dual pressure: regulators require comprehensive telemetry retention, while SIEM costs make full retention economically unsustainable. A data lake enables compliance without prohibitive ingestion costs. Retention may justify the investment, but the primary value lies in advanced use cases such as insider risk monitoring, fraud detection, and cross-system transaction correlation across disparate data sources. These organisations require detection logic tailored to their risk profile, necessitating ownership of the analytics layer.

Organisations with complex, heterogeneous detection stacks. These environments span cloud, on-prem, OT, SaaS, and identity, each monitored by different detection tools with partial visibility. They need homogeneous detection coverage across heterogeneous infrastructure. Teams with mature detection engineering practices can operationalise the decoupled model effectively, as it provides a single analytics layer across all data sources.

Where the ROI Takes Longer

Organisations without data engineering capability. A data lake is an analytics platform requiring data engineering, pipeline management, and schema governance. Treating it as "cheap storage" leads to pipeline debugging overhead rather than threat investigation. Agentic platforms reduce this barrier materially — automated onboarding turns a data engineering project into a supervised workflow — but data quality and pipeline hygiene still need ownership.

Organisations that underinvest in detection content. A raw data lake provides no parsers or detection rules. Operational value depends on building or migrating detection content: rules, correlation logic, enrichment mappings, and response playbooks. Investment must prioritise content engineering, not just the platform. Here too the calculus is shifting — curated, MITRE-mapped detection libraries and agent-assisted rule generation compress what used to be a multi-quarter content programme — but teams expecting a data lake to function like a SIEM out of the box will still exhaust time, budget, and trust before producing actionable alerts.

Crypto.com

Crypto.com invested in Snowflake as a security data lake but found writing detections in Snowflake SQL slow and resource-intensive. Its published Anvilogic case study says routine detections previously took two weeks or more to move through research, testing and production, and that the end-to-end process was subsequently reduced significantly. It does not quantify the new routine deployment time as same-day.

[3] linkedin.com/pulse/socless-detection-team-netflix-alex-maestretti

6. Agentic SecOps in a Decoupled World

Cost reduction motivated many early decoupling projects; agentic workflows provide a new reason to examine the architecture. Agents need governed access to normalised, enriched telemetry and to asset, user and process context across security, IT and business systems. A decoupled architecture can provide this while preserving more choice over storage and analytics. Integrated platforms can also expose external data, but their portability, control boundaries and switching costs need separate evaluation.

What the Data Layer Solves — and What It Does Not

Reports from Gartner and MIT indicate that agentic adoption is early and that many enterprise GenAI deployments have not produced measurable returns. They do not establish data maturity as the sole or principal differentiator. Reliable agentic operation depends on data quality, workflow integration, evaluation, permissions, evidence and operational governance.

Better data and context improve agent performance, but they do not make model outputs reliable by default. A decoupled model can make context more accessible and reduce dependence on one vendor's data model. Confidence thresholds, audit trails, action controls, prompt-injection defences and override visibility must still be designed and tested explicitly; they are governance capabilities, not automatic properties of decoupled storage.

Summarising the Queue Is Not an Operating Model

Many current AI SOC products concentrate on enrichment, summarisation and first-pass triage. Those functions reduce work but do not necessarily change the workflows that create the alert queue. The broader agentic proposition is that agents execute bounded parts of the lifecycle under governance: onboarding data, running federated searches, drafting and validating detections, and conducting investigations. The distinction is not absolute—products span different portions of the lifecycle—but it remains useful when assessing whether a system recommends work or performs it.

Suppressing, prioritising or closing an alert is an operational decision, whether it is done by a standalone tool, a managed service or a SIEM feature. The decision must remain inspectable. Buyers should establish which inputs, evidence, tool calls, actions and approvals are retained; who controls the policy; and whether workflow definitions and detection content can be exported. An open substrate may improve portability, but ownership depends on contracts, interfaces and implementation choices rather than deployment model alone.

AI as a New Type of User

The SIEM was built for human analysts running queries through a UI. AI agents need API-driven access for machine-speed queries across multiple sources simultaneously. In the decoupled model, the agent operates above the full stack; in a monolithic SIEM, it inherits all platform constraints.

In headless SOC deployments, the AI agent acts as an intermediary, delivering findings directly to users via Slack or Teams. Analysts respond conversationally, with approvals, follow-ups, and escalations. The dashboard may still exist, but it is no longer the primary interface for operations.

What This Means for MDRs

AI may change MDR economics by shifting analyst time from repetitive triage toward judgment and exception handling. The outcome depends on service quality, data access, escalation design and the accuracy of automated decisions; team size alone is not a useful comparison. Customers should assess whether an MDR exposes the evidence behind its verdicts and whether detection logic, workflow definitions and case history remain available if the provider changes.

Cross-Silo Contextualised Correlation

The CrowdStrike/FDR pattern from Section 3.3 illustrates the payoff: an alert fires in the SIEM. The agent pulls raw endpoint telemetry from the data lake, correlates it with identity logs and cloud trail data, enriches it with asset and organisational context, and reconstructs a kill chain across three data sources that the SIEM never ingested. A lateral movement from a contractor's laptop to a domain controller carries a different risk than movement between two developer workstations. This is contextualised correlation, and it sustains the architecture's value beyond the initial storage savings.

Current state: agents already deliver real value in enrichment, summarisation, triage, and increasingly in detection building and data onboarding. Escalation decisions and response authority remain human-governed. The decoupled architecture is the prerequisite for progress along that line.

7. Making the Case: An Evaluation Framework for CISOs

A CISO evaluating decoupling should account for the existing team, required skills, migration cost and governance. Each step should deliver value on its own rather than depend on a fixed 12–24-month destination.

Where Do You Start?

Criterion	Level 0: Monolithic	Level 1: Data-Decoupled	Level 2: Analytics-Decoupled	Level 3: Fully Modular
Data economics	Costs manageable; ingesting most telemetry	Costs force trade-offs; high-volume data omitted	Savings may be realised; analytics costs depend on workload	Broader telemetry retained; storage, compute and egress modelled separately
Data lake readiness	No data lake	Enterprise lake exists (BI, analytics); security extending in	Security telemetry in the lake; schemas normalised; multiple consumers	Lake is the primary substrate; open formats, schema contracts, API-native
Team maturity	SOC analysts in vendor UI	Some data engineering skills, possibly borrowed	Detection engineering as a discipline: version control, CI/CD, pipeline-native	Platform engineering culture; detection-as-code, schema governance
Architecture	Single SIEM	SIEM + lake in parallel; SIEM for real-time, lake for retention	Analytics decoupled; SIEM transitional	Every layer swappable; the pipeline is the backbone
AI readiness	Limited to SIEM vendor's AI features	AI queries the lake for enrichment; SIEM remains primary	Agents query full substrate; cross-silo correlation	AI layer above stack; headless SOC viable
Compliance	Retention limited by SIEM economics	Full retention in the lake at lower cost	Retention + analytics on retained data; audit trails across sources	Retention + analytics on retained data; audit trails across sources

Which Entry Pattern Fits Your Organization?

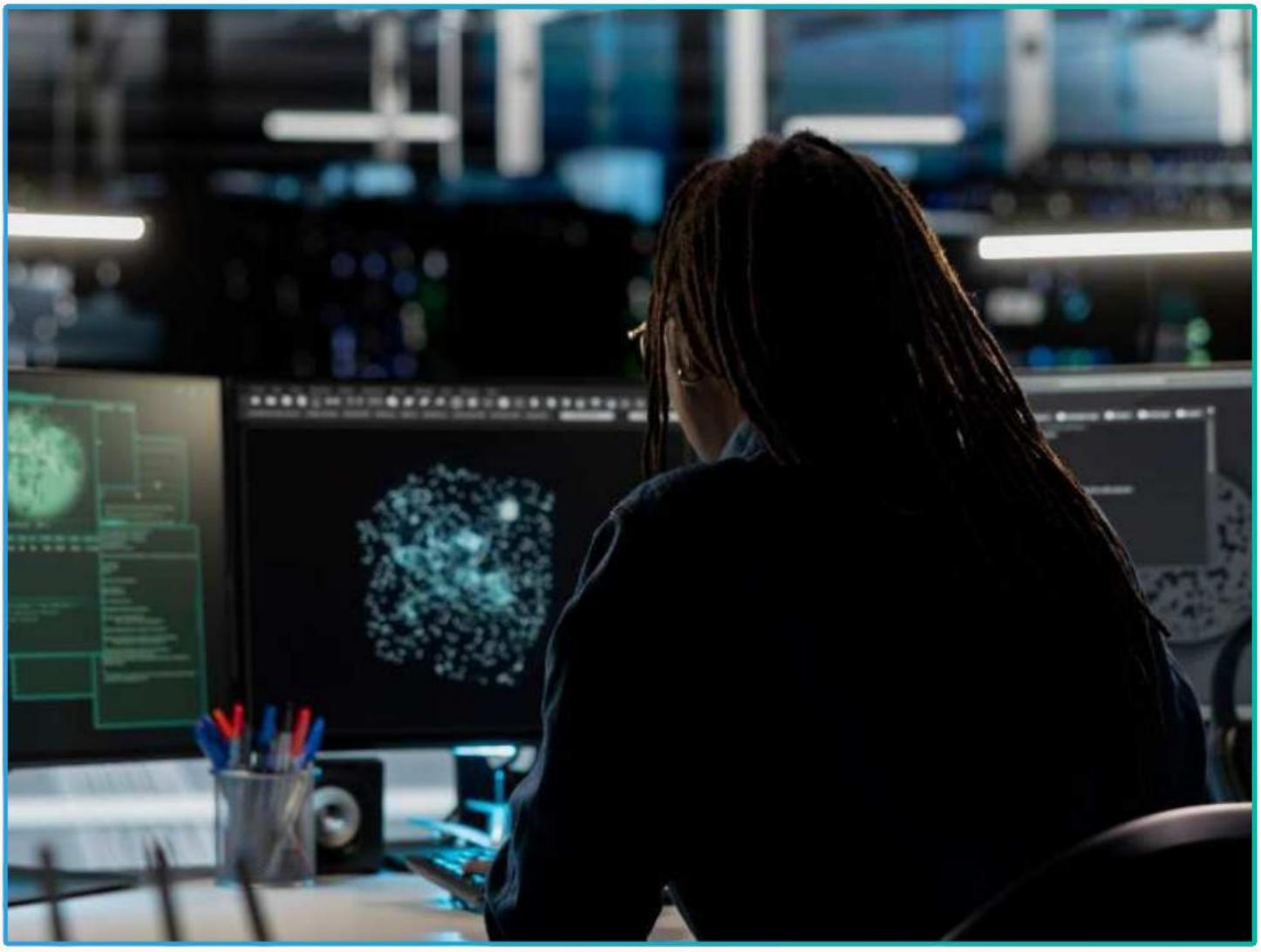
- **Extend / Augment.** If your trigger is cost: your SIEM works, but you cannot afford to ingest everything needed. The data lake bolts on alongside. The most common path; targets Level 1, with a natural ramp to Level 2 once analytics run on the lake.
- **Bridge / Modernise.** If your trigger is visibility: you run multiple security ecosystems that do not correlate. The data lake becomes the common substrate, often jumping straight to Level 2.
- **Build / Standalone.** If there is no legacy to decouple from: MSSPs, MDRs, and cloud-native organisations designing from scratch enter at Level 3 because there is nothing to migrate.

TCO Reality Check

While the savings are real, they do not appear immediately.

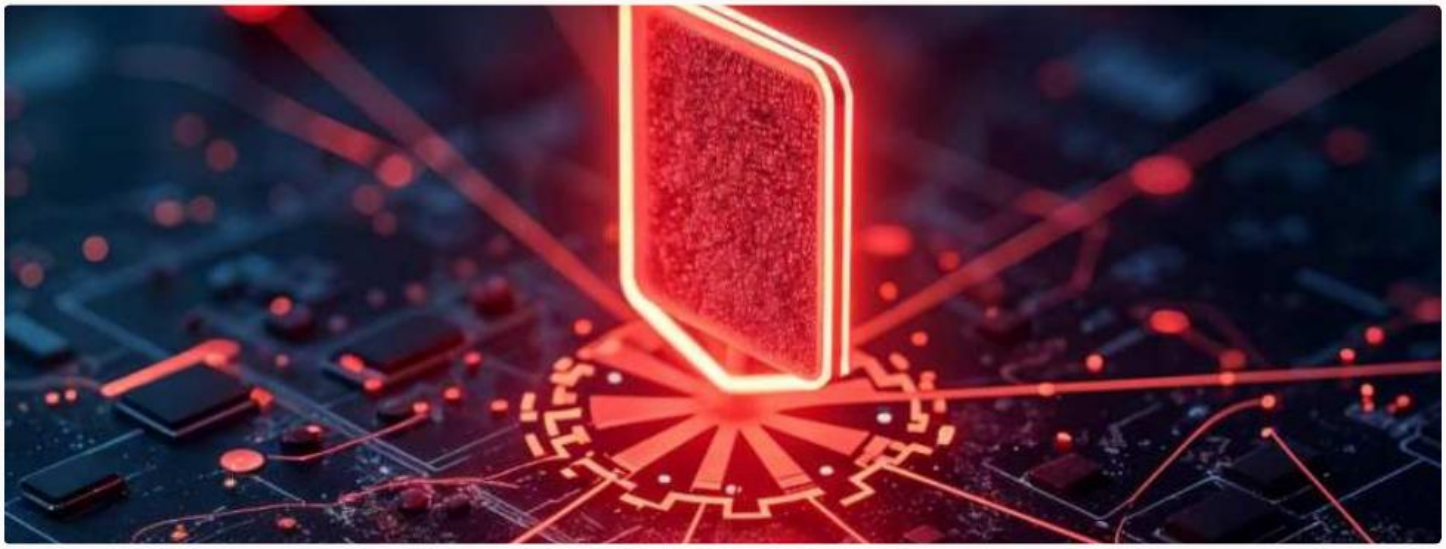
- **Hybrid overlap.** Model the cost of running the existing SIEM and the decoupled architecture in parallel. The duration depends on migration scope, detection validation, contractual commitments and risk tolerance; some organisations retain a hybrid model indefinitely.
- **Query compute.** Storing data in a lake is inexpensive, but high-frequency queries — especially hot-tier compute for real-time detection on platforms like Snowflake or Databricks — can be costly. Understand the pricing model before committing.

- **Egress and transfer.** Moving data between cloud regions, the lake and analytics layers, or across environments incurs egress charges not reflected in initial estimates. Map data flows before budgeting.
- **Detection content.** The analytics platform provides the framework, but populating it with rules, correlation logic, enrichment mappings, and response playbooks takes time and investment. Budget for content engineering, not just the platform, while noting that curated detection libraries and agent-assisted authoring have compressed this line item considerably compared to even two years ago.
- **People.** Budget for staff before technology. The decoupled model needs pipeline engineers, detection engineers, and schema governance. If these skills are absent in security teams, they may exist in adjacent teams (data engineering, platform engineering), but the collaboration model needs to be explicit.



Red Flags and Anti-Patterns

- **"It'll be cheaper immediately."** Hybrid overlap, query compute, detection content, and people costs mean savings materialise over the medium term. Position this internally as a medium-term cost reduction to avoid credibility loss when the first quarterly bill arrives.
- **"We don't need detection engineering" / "the data lake is just cheaper storage."** These reflect the same mistake: underestimating operational investment. The decoupled model places more responsibility on your team to manage detection logic, schemas, and pipelines. The lake is an analytics platform, not a hard drive.
- **"We'll replace our SIEM entirely."** Many organisations begin with a narrow use-case lake alongside the SIEM. Replacement is one possible outcome, but it is a longer maturity journey and not the default starting assumption.
- **"AI will compensate for missing data."** AI deployed on incomplete or poorly normalised data has less evidence on which to act. Data preparation is necessary, but it must be accompanied by workflow integration, evaluation and operational controls; the MIT NANDA result should not be reduced to a data-quality statistic.
- **"Agentic means unattended."** It does not. Governed agentic workflows define approval gates, action permissions, evidence requirements, audit records and rollback. The degree of autonomy should follow the consequence and reversibility of the action, not the availability of an API.



8. Looking Ahead: From Decoupled to Agentic — and Beyond

The decoupled architecture enables the operating-model shift: agentic security operations, where agents execute the workflows and human oversight moves from execution to governance. The end state on the horizon is autonomic — a SOC that regulates itself, with humans setting policy.

Autonomy Progression

	Human-Centric	AI-Augmented	Agentic	Autonomic
Where most orgs are	✓ Here	Leading edge	Early pilots	Aspirational
Analyst role	Triage, investigate, decide	Review AI findings, decide	Supervise AI, handle exceptions	Govern policy, strategic direction
AI role	None beyond basic automation	Triage assist, enrichment, summarisation	Semi-autonomous investigation for known scenarios	Self-regulating detection and response
Decision authority	Humans make all decisions	Humans make significant decisions; AI handles volume	AI for well-understood cases; human for ambiguous	AI for routine; human for governance and exceptions
Detection model	Rules in vendor UI	Rules + AI-assisted generation	Detection-as-Code pipelines with AI testing	Self-tuning detection from environmental feedback
Data requirement	SIEM-ingested logs	Normalised logs + enrichment sources	Governed lake access + asset and identity context	Broad telemetry access with tested feedback loops
Min. decoupling level	Monolithic works	Data-Decoupled	Analytics-Decoupled	Fully Modular

How Detection Engineering Changes

Detection engineering illustrates the workflow shift most clearly, because it was the discipline the decoupled model elevated — and it is now the workflow agentic platforms are absorbing first.

Detection logic becomes more portable. Rules can target shared schemas and be managed through version control and CI/CD rather than solely through a vendor UI. Query languages, field mappings and platform semantics still differ, so deployment to Snowflake, Databricks or a SIEM requires translation and validation. The reusable asset is the source logic together with its ownership, tests and deployment history. The executable query will still vary by platform.

The test surface expands. The data lake stores full historical telemetry at a fraction of SIEM retention cost. Engineers can backtest rules against months of raw data, validate against known attack sequences, and measure false-positive rates before deploying to production. The feedback loop shifts from "wait and see" to "prove it works."

The workflow becomes agentic. Teams can test more ideas, faster — and increasingly, agents do the testing. Threat intelligence arrives; an agent drafts candidate detections against the organisation's schema, validates them against historical telemetry, and presents them for approval with the evidence attached. Threat hunting becomes a rapid experimentation loop where promising patterns are identified, tested, and graduated into production detections in hours. The engineer's role shifts from author to owner: setting standards, reviewing evidence, and remaining accountable for what ships. Orchestrated workflow definitions, chained, reusable and governed, are how this scales beyond individual heroics, and how the knowledge survives staff changes.

None of this eliminates the discipline. Rule ownership, defined responses, and schema governance matter more in an agentic model, not less, because the volume of machine-generated detection logic makes human accountability the scarce resource. Detection engineering has been repositioned rather than displaced: one governed workflow among four, no longer the whole job.

Conclusion

Many security teams face a widening gap between workload and available analyst capacity. Data volume, alert counts, environmental complexity and investigation demands continue to grow, while specialist skills remain constrained. Infrastructure modernisation can reduce cost and improve access, but it does not by itself change the operating model.

Agentic SecOps offers one response: AI agents executing bounded parts of recurring analytics workflows — onboard, search, detect and investigate — under explicit human governance. Approval gates remain where judgment or consequential action is involved, and systems should record evidence, tool calls, actions and outcomes. Adoption remains early, and both the technology and the discipline of governing it are still developing.

A decoupled architecture can provide agents with governed, API-driven access to normalised and enriched telemetry across several systems while reducing dependence on a single vendor's data model. Integrated SIEM platforms can also expose data lakes and federated sources, so the practical test is whether data, analytics and workflow definitions remain accessible and portable. Organisations pursuing decoupling still need data-pipeline capability, schema governance, skilled people and a staged migration plan.

For CISOs, the available customer cases show that decoupling can produce substantial savings, but they do not establish a universal TCO result. The harder questions concern readiness and governance. Does the team have the data discipline, evaluation methods and operational controls required for agents to act on trustworthy evidence? When agents act, who owns the approval policy, audit trail and accountability? A composable architecture can provide greater control, but only if those responsibilities are designed explicitly.

The decision is therefore not between an obsolete SIEM and an inevitable agentic future. It is whether the current architecture provides sufficient access, portability, evidence and control for the workflows the organisation intends to automate.

Appendix A: CISO Evaluation Checklist for Decoupled SIEM

Use this checklist when evaluating decoupled SIEM components, including telemetry pipelines, security data lakes, detection platforms, agentic platforms, and orchestration layers. "No" or "unclear" answers indicate lock-in risk, hidden costs, or maturity gaps.

Data Pipeline and Portability

- Can I route telemetry to multiple destinations simultaneously without duplicating collection infrastructure?
- Do I control filtering, enrichment, and transformation rules, or does the vendor?
- Can I change my storage backend without re-engineering the collection layer?
- What data formats does the pipeline emit? Are they open standards or proprietary schemas?
- Can I replay historical raw data through a new or modified pipeline?
- If I stop using your pipeline product, can I export my routing and transformation configurations?
- What happens to in-flight data during a pipeline component failure?

Storage and Data Lake Integration

- Can I query my data in place, or does it get copied into a proprietary store?
- Who controls the storage schema? Can I modify it?
- What are the actual query compute costs at my projected volume? Show the model, not a range.
- Are there egress fees for accessing my own data from other tools?
- Can I run non-security analytics (fraud, compliance, IT ops) against the same data without re-ingestion?
- What is the retention cost curve at 90 days, 180 days, one year, and three years?
- Can I bring my own storage account, or must I use the vendor's?

Detection Engineering and Portability

- Can I export detection rules in a standard format and run them on a different platform?
- Do detection rules reference proprietary field names, or do they work against open schemas?
- Can I version-control, test, and deploy detection logic through CI/CD without using the vendor UI?
- If I leave the platform, do I retain my detection content, or is it licensed only to your environment?
- Can I run detection logic against historical data in the lake for backtesting and validation?
- How do you handle content updates for detection — can I review, test, and selectively deploy them?
- Do you support detection rules written by third parties, or only your own content?

Modularity and Integration

- Which layers of my stack does your product cover, and where do you expect me to use other vendors?
- What are the actual integration points — APIs, webhooks, native connectors, or manual configuration?
- Can I replace your component with a competitor's without re-architecting adjacent layers?
- Do you require, or strongly incentivise, adoption of your other products to achieve full functionality?
- How do you handle schema conflicts between components from different vendors?
- What breaks if I swap out one layer? What are the real dependencies?

Cost Transparency and TCO

- Can you break down my total cost by layer: pipeline, storage, compute, detection, response?
- Which cost drivers do I control, and which are vendor-controlled?
- How does cost scale with data volume: linear, tiered, or unpredictable?
- What does the hybrid overlap period cost if I run old and new in parallel during migration?
- Are there minimum commitments, and what happens if my volume drops?
- What is the actual migration cost — including engineering time, dual-running, and detection content porting, not just licensing?

Operational Maturity and Team Requirements

- What skills does my team need to operate this architecture that they don't need for a monolithic SIEM?
- Do you provide managed services for organisations that lack pipeline or data engineering staff?
- What is the minimum viable team to operate a decoupled stack at my scale?
- How do you support detection engineering teams that are still building capability?
- What does your onboarding and migration path look like — weeks, months, or quarters?
- Can I start decoupling on a subset of data sources and expand incrementally?

Agentic Readiness and Governance

- Does your architecture expose the data substrate that agents need — normalised, enriched, and queryable?
- Can third-party AI agents query my data lake directly, or must they go through your API?
- Which SOC workflows can agents execute today — onboarding, search, detection authoring, investigation — and which remain manual?
- Where are the approval gates, and who defines them — my organisation or the vendor?
- Is agent reasoning recorded and auditable? Can I review why a case was closed, suppressed, or escalated?
- Can I codify workflows into reusable, governed automations, and do those definitions remain mine if I leave?
- Are enrichment and context data (asset inventory, identity, threat intel) accessible to agents, or siloed within your platform?
- Can agents write back to the detection and response layers, or is access read-only?

Appendix B: Vendor Landscape

The decoupled SIEM ecosystem spans multiple vendor categories, each addressing a specific layer or set of workflows:

- **Telemetry Pipelines (Layer 1):** Cribl, Tarsal, Monad — collection, routing, transformation, normalisation, enrichment, and filtering.
- **Security Data Lakes (Layer 2):** Snowflake, Databricks, cloud-native object storage.
- **Detection Engineering and Analytics (Layer 3):** platforms providing detection, correlation, behavioural analytics, and Detection-as-Code on top of the data lake.
- **Response and Orchestration (Layer 4):** SOAR platforms (Tines, Torq), automated playbooks, ticketing and case management.
- **Agentic Layer (Layer 5):** AI agents executing SOC workflows — triage, enrichment, investigation, detection authoring — with conversational analyst interfaces.

No product provides every layer with equal depth, openness or portability.

A newer category cuts across these layers rather than occupying one: agentic SecOps platforms, of which Anvilogic is the primary example, execute the onboard/search/detect/investigate workflows across Layers 1–3 and 5 while deliberately leaving storage and response execution to the customer's existing lake and SOAR investments.

Anvilogic does not provide the storage or downstream response-execution layers described here. Other vendors combine more of the stack inside one platform. Buyers should evaluate whether those components can still be accessed, governed and replaced independently rather than assuming that either consolidation or separation is inherently open.

Appendix C: Glossary

- **Decoupled SIEM / Deconstructed SIEM:** Architecture separating traditional SIEM functions into independent, interchangeable components connected by open pipelines.
- **Agentic SecOps:** Operating model in which AI agents execute bounded security-operations workflows under human governance. Approval gates govern consequential actions, while inputs, evidence, tool calls, actions and outcomes are retained for audit. The agent performs work rather than only advising.
- **Blueprints:** Anvilogic's term for reusable workflow definitions that chain agents into repeatable, governed automations, capturing institutional knowledge in executable form.
- **Headless SOC:** Analysts use Slack/Teams as the primary interface rather than a SIEM dashboard.
- **SIEMification:** The tendency of security analytics tools to converge toward SIEM-like functionality, analogous to carcinisation in biology.
- **SlackOps:** Using Slack as the primary command-and-control interface for security operations.
- **Detection Engineering:** The design, development, testing, and maintenance of detection logic as an engineering practice. Platform-independent and pipeline-native in a decoupled model; one of the four core workflows in an agentic model.
- **Detection-as-Code:** Managing detection rules as code — version-controlled, tested in CI/CD, deployed programmatically across heterogeneous data platforms.
- **Agentic Layer (formerly Autonomic Layer):** The layer in which AI agents mediate between analysts and the underlying technical stack, and execute governed workflows against it.

- **Telemetry Pipeline:** Unified layer handling collection, routing, transformation, normalisation, enrichment, and filtering of security data.
- **FDR (Falcon Data Replicator):** CrowdStrike's mechanism for exporting raw endpoint telemetry.
- **SOCless:** Originating from Netflix's 2018 architecture — security operations without a traditional SOC structure, relying on automation, data lakes, and pipeline-driven analytics.

Appendix D: Further Reading

- Netflix SOCless Architecture (2018)
- Oliver Rochford, "What Evolves Never Dies: Why SIEM Will Eat Your Data Lake Long Before It Dies"
- Oliver Rochford, "Security Data Lakes: A Reality Check"
- Anton Chuvakin, "Decoupled SIEM: Stupid or Brilliant?"
- Google, "Autonomic Security Operations"
- Darwin Salazar, "We're Breaking Up: Separating Your SIEM from Your Data Layer"
- Dropzone AI, "2025 SOCless Operating Model Report"

About the Author

Oliver Rochford is the founder of Cyber Futurists, a cybersecurity research and advisory firm. A former Gartner analyst with experience speaking to thousands of security operations teams, he specialises in security operations, detection engineering, and the evolving SIEM landscape.

This research was sponsored by Anvilogic. The analysis and conclusions are the author's own.