SACR

NOVEMBER
2025

REALM.
SECURITY

THE RISE OF SECURITY DATA PIPELINE PLATFORMS AS A CONTROL PLANE FOR THE SOC

AQSA TAYLOR
CHI AGHAIZU

Research

We explore the newest frontiers of cybersecurity.

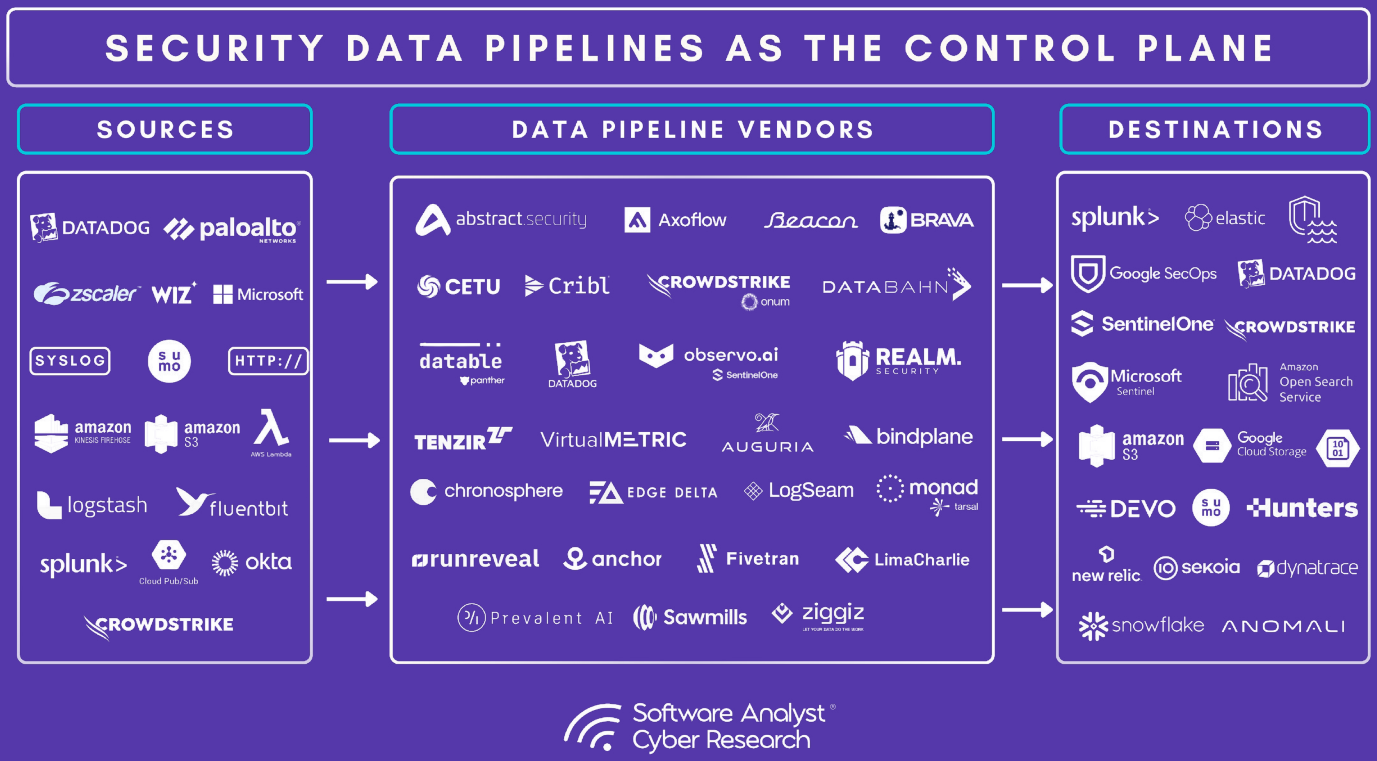
Whether you're looking at emerging vendors, evolving threats, or shifting architectures, our timely, opinionated insights help modern security leaders make smarter, faster decisions.

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.



Software Analyst Cyber Research

Author

- **Aqsa Taylor** Chief Research Officer, leading research division, CISO arm and AI product development at SACR. She is a published author of two cybersecurity books and numerous research articles on cloud security and SecOps. She brings a decade of experience as a Product Leader with a track record of building some of the well known security platforms such as Twistlock, Prisma Cloud Workload protection, Prisma Cloud Agentless scanning, Gutsy (Minimus), merging cybersecurity, AI, and cloud infrastructure. With deep experience across early-stage startups, she has led stealth to launch product efforts, integrated multi-million dollar acquisitions, and delivered features that drive real value for users. She is also a public speaker with writings recognized across several renown media outlets.

Co-Author

- **Chi Aghaizu** is an award winner Engineer and brings four years of specialized experience building and securing AI systems, with expertise in LLM integration, RAG, and MLOps practices focused on data quality, governance, and auditability. Her background in data pipeline architecture and transformation workflows positions her to analyze how security data flows through enterprise infrastructures, identifying vulnerabilities in collection, normalization, and routing mechanisms. As an award-winning developer, Microsoft Certified Azure AI Engineer, and holder of dual Master's degrees, she combines hands-on engineering experience with analytical rigor to assess enterprise security solutions against real-world operational demands.

Quick Read:

Here's quick insights on the report

1. Market Consolidation Accelerates with Acquisitions

The SDPP market is entering a rapid consolidation phase as major SIEM, XDR, and observability providers acquire pipeline platforms to strengthen their data architectures. Recent deals include CrowdStrike acquiring Onum for about 290 million dollars, SentinelOne acquiring Observo AI for approximately 225 million dollars, Panther Labs acquiring Datable for an undisclosed amount, and now, Palo Alto Networks announced acquisition of Chronosphere, for $3.3B dollars (one of the biggest in the industry). These acquisitions reflect a clear industry trend. Large security and observability vendors are absorbing pipeline capabilities to overcome long-standing ingestion, normalization, and cost challenges within their own platforms. Buying is proving faster and more strategic than building, and this shift is moving the center of gravity in the SOC toward the pipeline layer.

But what does this mean for the vendor neutrality benefit that security data pipelines have long been known for? Practitioners express concerns about vendor neutrality, migration bottlenecks, and more in this report.

2. Security Data Pipelines Have Become the SOC Control Plane

Pipelines no longer simply move logs. They now govern ingestion, normalization, enrichment, routing, tiering, and data health. As a result, they have become the primary control plane of the modern SOC. Every downstream system relies on them for clean, consistent, and trustworthy telemetry. In this report, you'll find an indepth valuation of core capabilities and emerging innovation in the security data pipeline market.

3. AI Is Becoming Essential for Pipeline Operations

AI adoption is practical, assistive, and explainable. Security teams want AI that handles engineering-heavy or repetitive work such as parser creation, schema drift correction, pipeline generation, baselining, and anomaly detection. Teams aren't comfortable yet with autonomous decision-making in the SOC but strongly support AI within pipelines to reduce workload and increase consistency in pipeline operations.

4. Telemetry Health Monitoring Is Now Critical

Security teams express more fear of missing data than of noisy data. Pipelines now provide intelligent, continuous telemetry health: silent source detection, schema drift, volume anomalies, baseline deviation, noisy source spikes, and rerouting options based on destination failures. This monitoring in the data layer ensures the SOC never operates blind.

5. Shift Detections Left

Some platforms are pushing detections into the pipeline, performing lightweight IOC checks and early pattern recognition before events reach the SIEM. Practitioners value earlier context but note that response speed, not detection timing, often limits real impact. Learn more about how this trend is shaping impressions within the security community.

6. Pipelines Form the Foundation for AI Driven Security Operations

AI systems depend on high-quality, normalized, enriched, and complete data. Pipelines are becoming the preparation layer for AI copilots, LLM-based SOC assistants, advanced correlation engines, and autonomous triage. Without pipelines, AI performance degrades significantly. This makes SDPPs a strategic enabler for future SOC automation.

7. SDP PLUS Vision

In addition to core pipeline features, we are seeing a trend in which pure-play SDP platforms aim to expand horizontally across the SOC stack by taking on adjacent category capabilities beyond traditional pipeline functions. These include in-house data lake options with tiered storage, threat detection and analytics at the pipeline layer, federated search and querying across SIEMs and data lakes, observability convergence, and AI SOC-like capabilities.

Summary for Security Leaders

Security data pipelines are now the control plane of the modern SOC. They own data and deliver cost efficiency, improved data quality, faster investigations, cleaner enrichment, better telemetry reliability, and vendor-neutral routing. They are also becoming the data foundation needed for next-generation AI-driven operations. As acquisitions accelerate, the market is shifting toward two branches: standalone security data pipeline platforms and SDP capabilities within broader architectures. Either direction underscores the importance of the data pipeline as the most critical layer in the security stack.

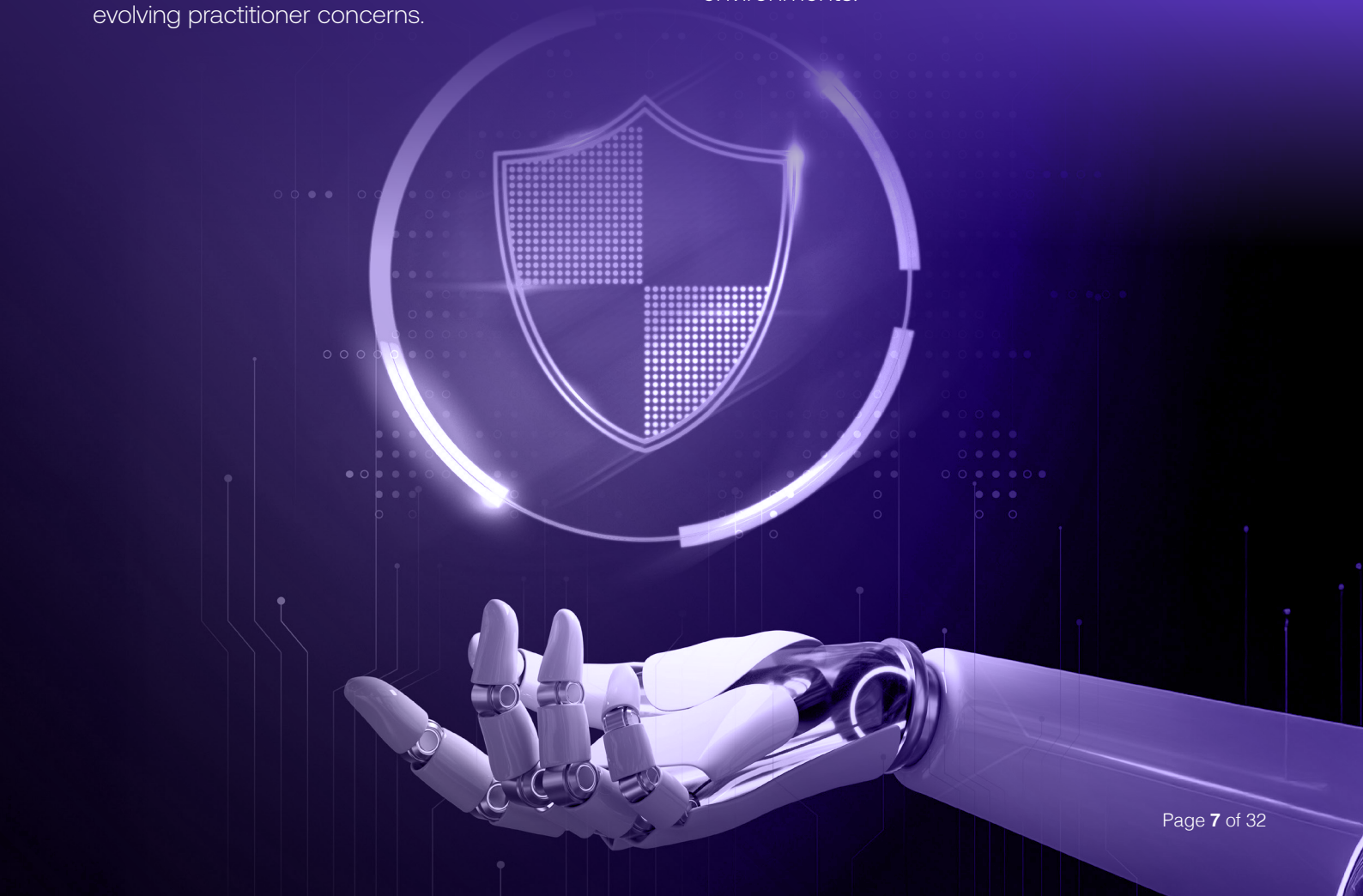
Table of Contents

Introduction

Before diving into this report, it's important to set some context for readers who are learning about security data pipeline platforms (abbreviated as SDPP throughout the report). In Francis's first report, ***The Market Guide 2025: The Rise of Security Data Pipeline Platforms***, he introduced what these platforms bring to the world of security operations. It was the first analyst report focused on this category, even though the solution had existed for years. Their rise was largely driven by practitioner concerns about legacy SIEM platforms and ongoing issues with data quality. I explained the practitioner concerns in detail in my ***Convergence of SIEM Platforms*** report, where I also highlighted how two major SIEM vendors acquired security data pipeline companies to redefine SIEM capabilities through pipeline integration. Since then, there has been another acquisition by Panther of the Datable.io security pipeline platform. And as of today, November 19th, Palo Alto Networks announces acquisition of Chronosphere.I expect we'll continue to see more of these acquisitions soon as SIEM vendors race to outpace legacy limitations and evolving practitioner concerns.

In this version of the report, we take a deep dive into the world of security data pipeline platforms, exploring how the category has evolved over the past year and the different directions vendors are now taking. The focus is on mapping how these platforms have matured in both capability and purpose, moving from basic data routing tools to core components of modern security architectures. This report expands on the original framework with a new layer of analysis that explores several key pipeline capabilities in depth, taking into account both the breadth and the maturity of the features.

The report further captures the different paths SDP vendors are taking, from those deepening their integrations with SIEMs as data routing platforms to those moving toward in-house data lake capabilities in a vision to rise as "SDP PLUS" platforms. It draws from practitioner conversations, customer interviews, and in-depth briefings to provide a grounded view of how these platforms are being adopted and adapted within real SOC environments.

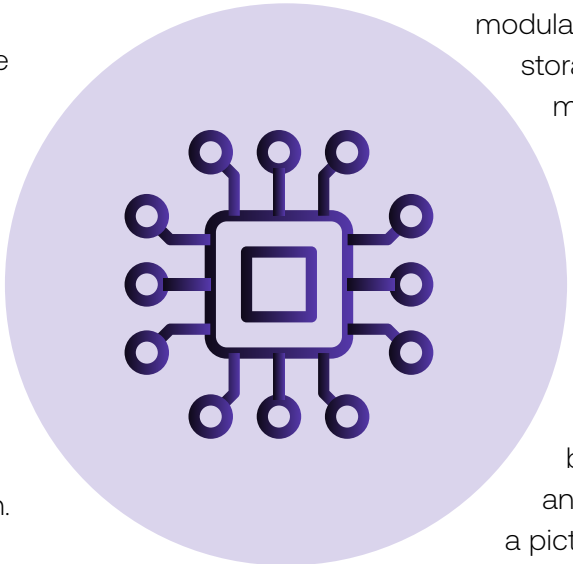


Revisiting the SDPP Report v.1: How the Data Layer Became the Heart of the Modern SOC

In our first report, we defined what Security Data Pipeline Platforms are -

Security Data Pipeline Platforms (SDPP) are purpose-built systems that ingest, normalize, enrich, filter, and route large volumes of security telemetry across hybrid and cloud environments. These platforms sit between data sources (like EDRs, cloud logs, and firewalls) and destinations (like SIEMs, data lakes, XDRs, and analytics tools). Their goal is to optimize the flow and quality of telemetry data to reduce operational complexity and cost while increasing the speed and accuracy of detection and response.

When the Security Data Pipeline Platform (SDPP) report was first published, it drew attention to something that many security teams had quietly been feeling for years. SIEMs were reaching a breaking point, yet, not disappearing. As organizations collected more data, the traditional model of ingesting everything was becoming impossible to sustain. The report highlighted that this shift marked a deeper transformation in how modern Security Operations Centers (SOCs) would be built with the introduction of security data pipeline platforms in the data fabric.



At the center of the report's findings was the rise of the SDPP. These platforms were described as a new foundational layer in the SOC, sitting between data sources and destinations like SIEMs, data lakes, and XDR tools. We called them the "security refinery" of the modern era because they clean, enrich, and route raw telemetry into structured, high-quality data that analysts can actually use.

It also made an important point about why this market was growing so fast. Data growth, rising compliance demands, and tool sprawl were all putting pressure on SOCs to find a more efficient way to manage telemetry. The report highlighted that SDPPs not only reduce cost but also improve the quality of data, helping faster threat detection.

We also highlighted how SIEM was evolving. Instead of serving as a single, monolithic system, the modern SIEM is shifting toward a modular architecture that separates storage from analytics. This new model allows data to live in cheaper cloud storage while being queried on demand, giving organizations the flexibility to scale without breaking their budgets.

The report clearly anticipated the convergence between pipelines, data lakes, and SIEM systems. It painted a picture of a security data fabric where ingestion, storage, and analysis would become part of one unified layer. For many in the industry, that idea shifted the conversation away from which SIEM to buy toward how to build the data architecture that supports it.

In short, Security Data Pipeline Platforms are becoming a must-have for modern organizations because they completely change how security data is collected, processed, and used. Here's a deep dive into why they matter and how they have now evolved to become the control plane for SOC.

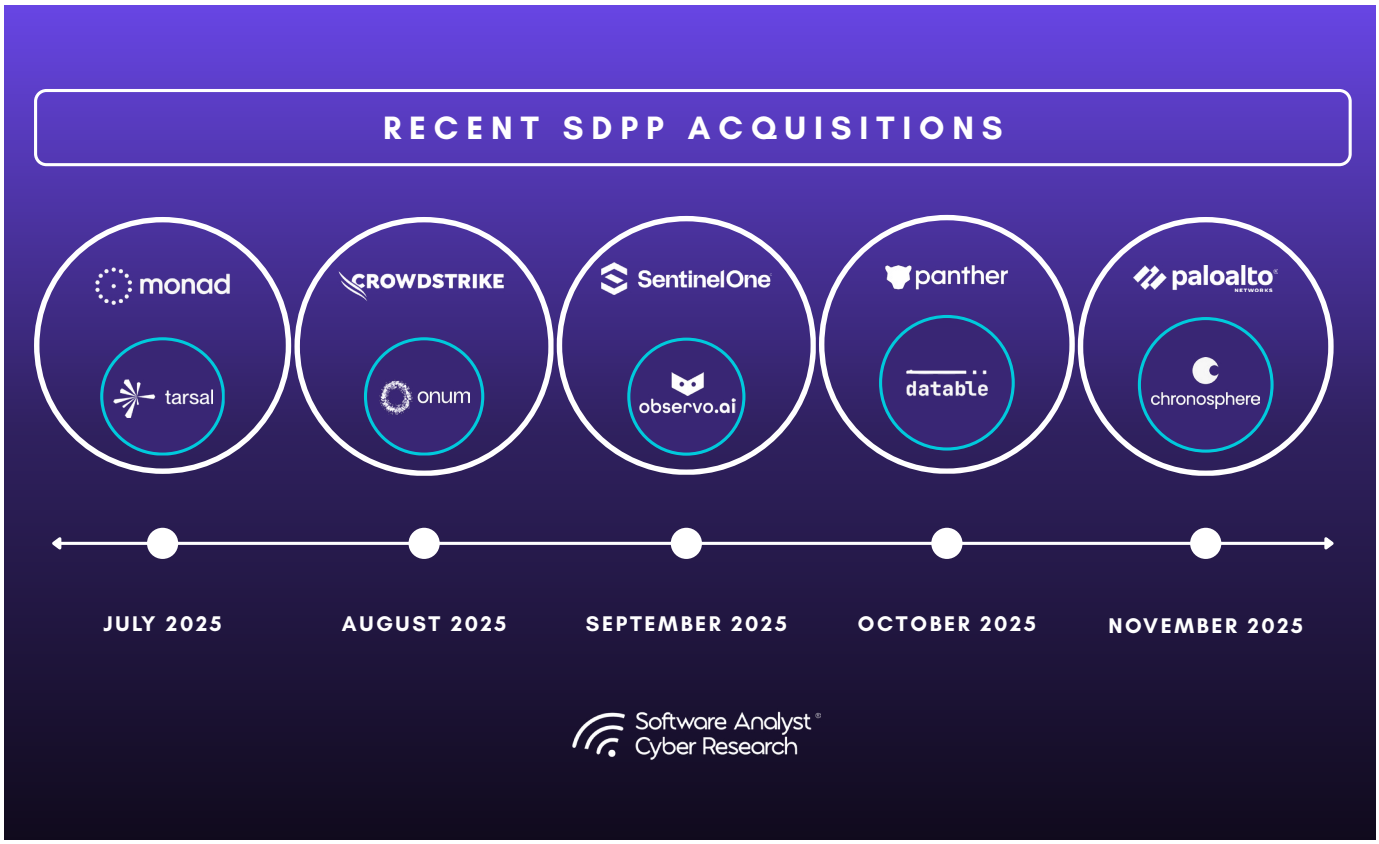
Acquisitions:
A Wave of Consolidation Across SDPP Vendors

Good security depends on good data. Because of this, SIEM and XDR vendors are moving quickly to control the pipelines that clean, shape, and route telemetry. This shift marks the beginning of a new phase in the market, where data quality becomes just as important as detection or response.

Over the past two years, several major security and observability companies have acquired smaller pipeline and telemetry vendors. This trend shows that the industry now understands how important the data pipeline has become. Modern security platforms need high quality, well prepared data before they can deliver strong analytics or AI driven outcomes. This growing importance is evident in the push from large vendors to bring pipeline technology in house instead of relying on third parties.

In chronological order of announcements -

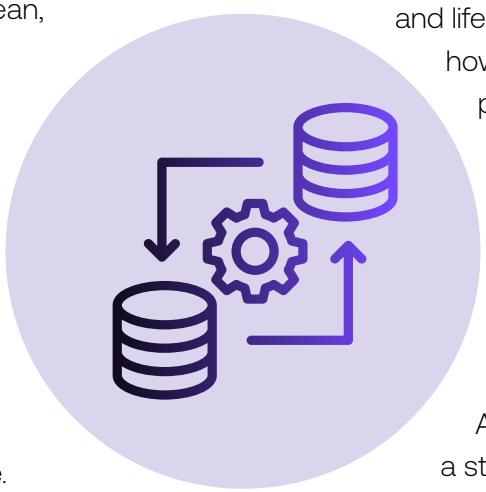
- Tarsal: Tarsal was acquired by Monad to enhance its security operations and data management capabilities in July 2025 - one of the first acquisitions.
- Onum: CrowdStrike acquired Onum for **about US$290 million.**
- Observo AI: SentinelOne acquired Observo AI for **approximately US$225 million** (cash + stock) to enhance its data-pipeline/SIEM capabilities.
- Datable: Panther Labs announced acquiring Datable, a security-data-pipeline platform. (Amount undisclosed)
- Chronosphere: Palo Alto Networks announced acquisition of Chronosphere, a major observability platform with pipeline capabilities for $3.3B dollars, on November 19, 2025.



What Acquisitions Mean for Broader Security Platforms

These acquisitions show a clear trend: Security Data Pipelines are becoming the control plane of modern security operations. Vendors want to sit closer to the source of data because strong AI and strong analytics depend on clean, well-structured telemetry. Instead of competing on dashboards or detection content, companies now compete on data quality, consistency, and readiness.

This shift brings clear benefits for customers of the broader platforms - the SIEMs. Performance improves, noise decreases, and storage costs go down. The core message is simple. Whoever owns the data quality and routing, has a larger play in the modern, decoupled, SOC architecture. And hence, the pipeline layer is becoming the heart of the SOC and the operational place where teams decide what data matters, how it should be shaped, and where it should go. It governs data quality, routing, enrichment, and lifecycle management, shaping how downstream tools perform.The platforms that integrate natively with Security Data Pipeline platforms will define the next generation of modern analytics platforms.



These acquisitions confirm that the future of SIEM, XDR, and AI SOC technologies will rely on a strong, unified control plane built in the pipeline layer. Whoever controls this layer ultimately controls the quality, cost, and intelligence of the entire SOC stack.

Neutrality Concerns with Acquisitions

Although the acquisition strengthens the larger platforms, it raises concerns about neutrality for users of the security data pipeline platforms.

As more SDPP vendors are acquired by large SIEM, XDR, and observability platforms, security leaders are beginning to express clear concerns. The biggest worry is the potential loss of neutrality. Many organizations adopted independent pipeline platforms because they provided flexibility, transparent routing, and the freedom to choose or change destinations without friction. When these platforms become part of a larger ecosystem, their priorities may shift toward favoring the parent vendor's integrations while downplaying independent capabilities. This can limit multi-destination routing, reduce portability, and recreate the very vendor lock-in that SDPPs were designed to eliminate.

There is also apprehension that innovation in the category may slow as acquired platforms are folded into broader product roadmaps. Independent SDPPs often moved quickly, responding directly to practitioner needs. Once inside a major vendor, development may be shaped by platform alignment rather than customer choice. For now, many of the acquired companies have shared that their vision is to continue supporting the standalone security data pipeline platform to its users without forcing lock-in. Whether this trend will continue to evolve, is to be determined.

The Evolution of Security Data Pipeline Platforms as a Distinct Category

Pipeline capabilities are sometimes merged into broader platforms such as SIEMs, observability tools, or XDRs. However, the focus of this report is primarily on what we refer to as “pure play” security data pipeline platforms.

Pure Play Security Data Pipeline Platforms

These platforms focus primarily on the data transformation layer between data sources and data destinations. We will see in the latter parts of the report, a trend where these platforms envision to take more of the adjacent capabilities gradually, we call that “SDP PLUS”, but in their current state, they still heavily fall under “Pure play” Security Data Pipeline Platforms.

Cribl (2018) stands as dominant Series E player with funding above **$600M**, and at $3.5B evaluation, embodying a broader shift from ***log routing*** to ***full security data pipeline platform.*** Cribl still stands at the center of the **Security Data Pipeline Platform (SDPP)** market as its most mature and influential leader both technically and commercially. Many of the practitioners we interviewed know the SDP market by Cribl’s name.

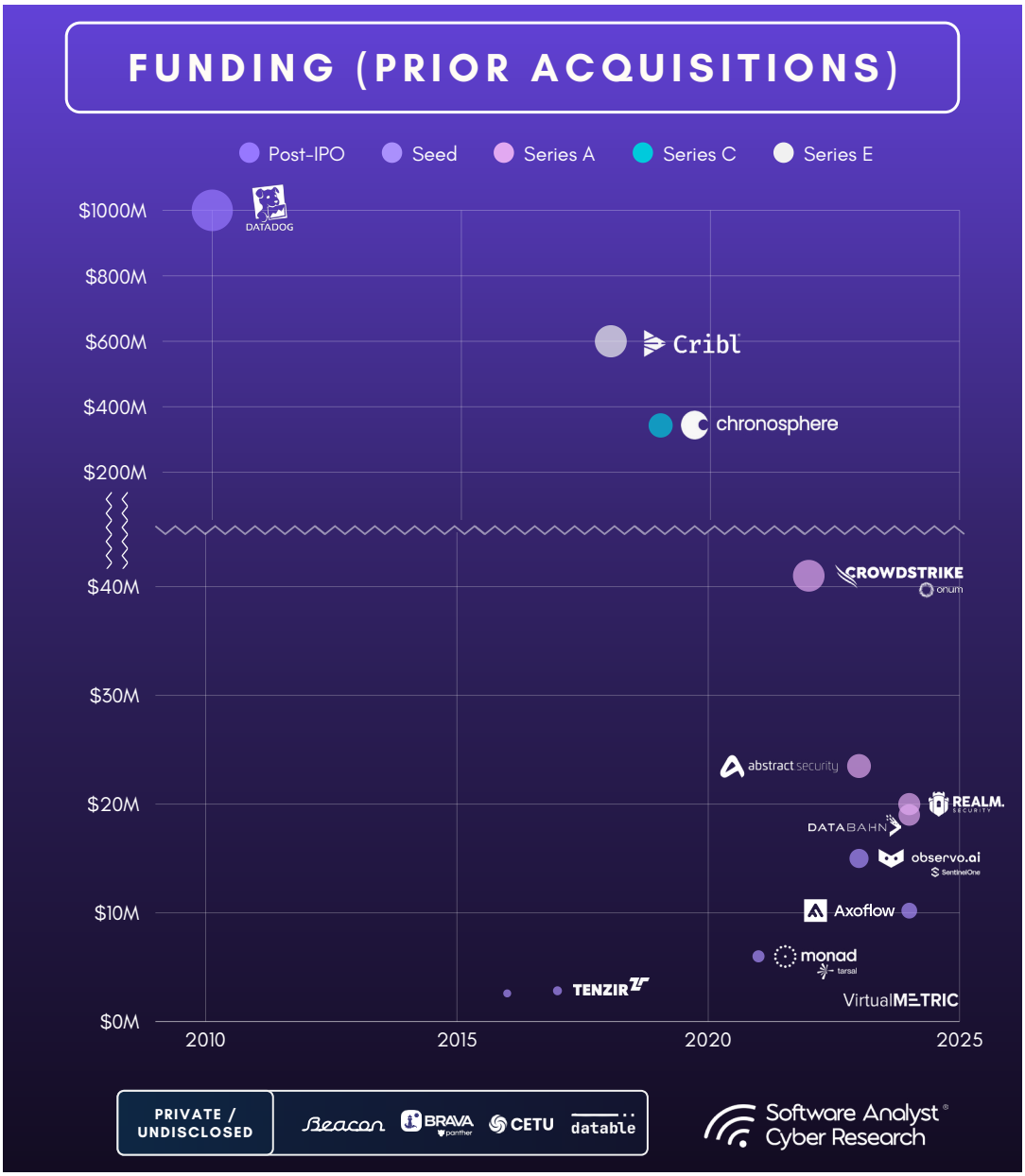
Emerging Entrants

In addition to Cribl, we’ve done an in-depth analysis of these emerging security data pipeline platforms in this report. In Alphabetical Order -

- Abstract Security founded in 2023, raised $15M dollars in Series A in 2024.
- Axoflow founded in 2023, seed round of $7M dollars in Jan, 2025.
- Beacon Security came out of Stealth in November, 2025.
- Brava Security - currently in stealth
- CeTu founded in 2024
- Databahn in 2023, $17M dollars in series A, in Jan 2025
- Datadog launches Observability Pipelines - June 2022
- Datable founded in 2023
- Onum founded in 2022, acquired by Crowdstrike
- Observo AI in 2022, acquired by SentinelOne
- Tenzir, founded in 2017, raised $3.3M dollars in seed round
- Realm Security founded in 2024, raised $15M in series A
- VirtualMetric, founded in 2025, raised $2.59M in seed round

While early infrastructure players consolidated around observability and log management, the platforms are now focusing on security and attacking ***specific SOC pain points***: data quality for security, ingestion cost, AI normalization, and cross-platform routing.

Growing Investment in the Security Data Pipeline Category



Insights

- The **average jump** from Seed to Series A/D rounds across these vendors ranges from **4x to 10x** in valuation, signaling strong investor confidence in SDPP platforms.
- Consolidation is accelerating : **Observo AI, Datable, and Onum** were acquired by major security players (SentinelOne, Panther Labs, and CrowdStrike).
- Palo Alto Networks announces acquisition of Chronosphere for $3.3B on Nov 19, 2025 .

Over the next three years, we anticipate capital will continue moving toward pipeline-driven ecosystems that combine telemetry management, AI readiness, and cost efficiency, forming the backbone of the next security data economy.

Security Leaders Voice

Now that we've recapped what happened during our last research, it's time to mention what we learned these past months from many practitioner calls, security vendor in-depth interviews and product briefing and a questionnaire that covered every single detail of platform capabilities.

Security data pipelines began as a cost-saving broker. But they are now strategic policy engines for visibility, control, and agility. The practitioners adopting these platforms are not chasing hype around autonomous SOCs. They are building disciplined, deterministic systems supported by selective automation. The future of detection will belong to teams that control their data with the same rigor they apply to threat response and SDPPs are becoming an important layer that make this possible. Data control is the new detection.

Across industries, from financial institutions to managed service providers and industrial operations, the message is consistent. Security data pipelines are no longer considered back-end utilities. They are becoming the operational control plane for telemetry, cost management, and detection agility.

Practitioners entered this space to reduce log costs, but they stayed because of control. By centralizing routing, transformation, and lifecycle management, the pipeline has shifted from infrastructure to intelligence. One leader summarized it plainly: "We are not just compressing data anymore. We are deciding what matters and where it should live."

These conversations show a shift from tool-centric thinking to outcome-centric design. Practitioners prioritize three things above all:

- The ability to reduce data ingestion at SIEMs and automatically express transformations
- Multi-tier intelligent routing that aligns storage cost with data purpose
- Built-in observability that measures ingestion completeness and source health

Ease of management is becoming a key differentiator. Teams managing multiple customer environments (MSSPs) prefer centralized templates where one pipeline update can propagate across tenants. Smaller organizations prioritize alignment with their deployment style, especially infrastructure-as-code.

Expanding on these use cases, we asked practitioners to stack rank SDPP capabilities. And here's what we found -

Budget Management as an Entry Point

The original motivation was budget pressure and it remains one of the biggest reasons. Teams set reduction goals without losing context, cutting ingestion volume while improving fidelity. One leader cited processing over three terabytes of daily data but forwarding less than half of that after filtering. The immediate benefit from SDP platforms is lower cost at destination, but the deeper change is operational freedom. In the words of one leader, "You cannot automate nonsense." Poor data quality is still the most expensive problem in the SOC.

Normalization is the New Norm

Every practitioner began their modernization story with data normalization. They view consistent schemas as the precondition for any mature detection or analytics program. When normalization is right at the start, vendor content and correlation logic across the stack finally work as designed.

Intelligent Pipelines

Modern designs favor data adjacency rather than consolidation. Practitioners anticipate and prefer Pipeline platforms to adopt AI capabilities or intelligent routing with an understanding of data to direct data to the most cost-effective and policy-compliant storage, whether local, cloud, or cold archival, without binding analysis to a single ecosystem. These leaders want to govern the full data lifecycle, deciding what stays hot, what rolls warm, and what archives cold with clear rehydration paths when investigations begin.

Noise is not the enemy, silence is

A recurring theme across interviews was the danger of quiet systems. Practitioners worry more about missing telemetry than excessive alerts. They described the problem of dormant integrations, acquisitions without visibility, and logs that silently stop forwarding. The emerging use case is "silence detection," where the pipeline monitors the health of every data source and flags anomalies in activity levels or schema freshness.

AI is Becoming Familiar

Industry is becoming more and more comfortable now with the idea of agentic AI or copilot capabilities, however, not really buying the "autonomous" messaging yet. Leaders emphasized they are not ready to hand decisions to autonomous agents. They do, however, welcome targeted automation that eliminates repetitive work. They want AI to generate parsers when formats change, to detect version drift, to cluster similar events, and to perform quality assurance on closed investigations. They want explainable automation, not invisible reasoning. The ideal is agentic assistance, not autonomous control - yet.

Shifting Detections Left, into the Stream

The idea behind this direction that a few security data pipeline platforms are indulging in, is to detect threats based on IOCs while data is streaming through your security data pipeline. By moving detections into the stream and closer to the pipeline, you bring detection logic closer to the source and avoid the post-index costs or latencies that occur at SIEM destinations. This results in faster threat detection and reduces MTTD with near real-time speed.

While the concept in theory sounds impressive, in our interviews with practitioners, it received mixed feedback. Some welcomed the visibility into threat earlier in the stream, but some suggested they didn't see speed of detection in stream as their priority when speed of remediation is yet to catch up. Among those who saw value, organizations are experimenting with lightweight detection logic in stream, with an aim to add more context to the data that is routed to destinations. The goal is not to replace centralized analytics but to reduce dwell time and stage response earlier. Several teams already use the pipeline to automatically collect forensics when certain triggers appear, with detection in stream, the idea is to now surface these triggers without post-index delays saving time on detection and latency.

Acquisitions and the Question of Losing Neutrality

Early adopters embraced SDPs because they sat between systems and provided architectural control, flexibility, and cost savings without locking the customer into a single platform. That neutrality was the differentiator. Now, as major SIEM and data infrastructure players acquire pipeline companies or replicate their features, the market risks returning to the very vendor dependency that SDPs were meant to eliminate.

In the practitioner's words, "we're just going to end up back where we started, everything re-bundled under one large platform." The most valuable providers will integrate broadly across SIEM, observability, and data lake layers while keeping control in the hands of practitioners. The differentiator will not be who owns the data, but who enables transparent, vendor-agnostic flow across it.

If pipeline vendors continue to prioritize openness and integration, they can remain the connective tissue of modern security architectures. If they instead chase full-stack ownership, they risk becoming another feature in someone else's platform.

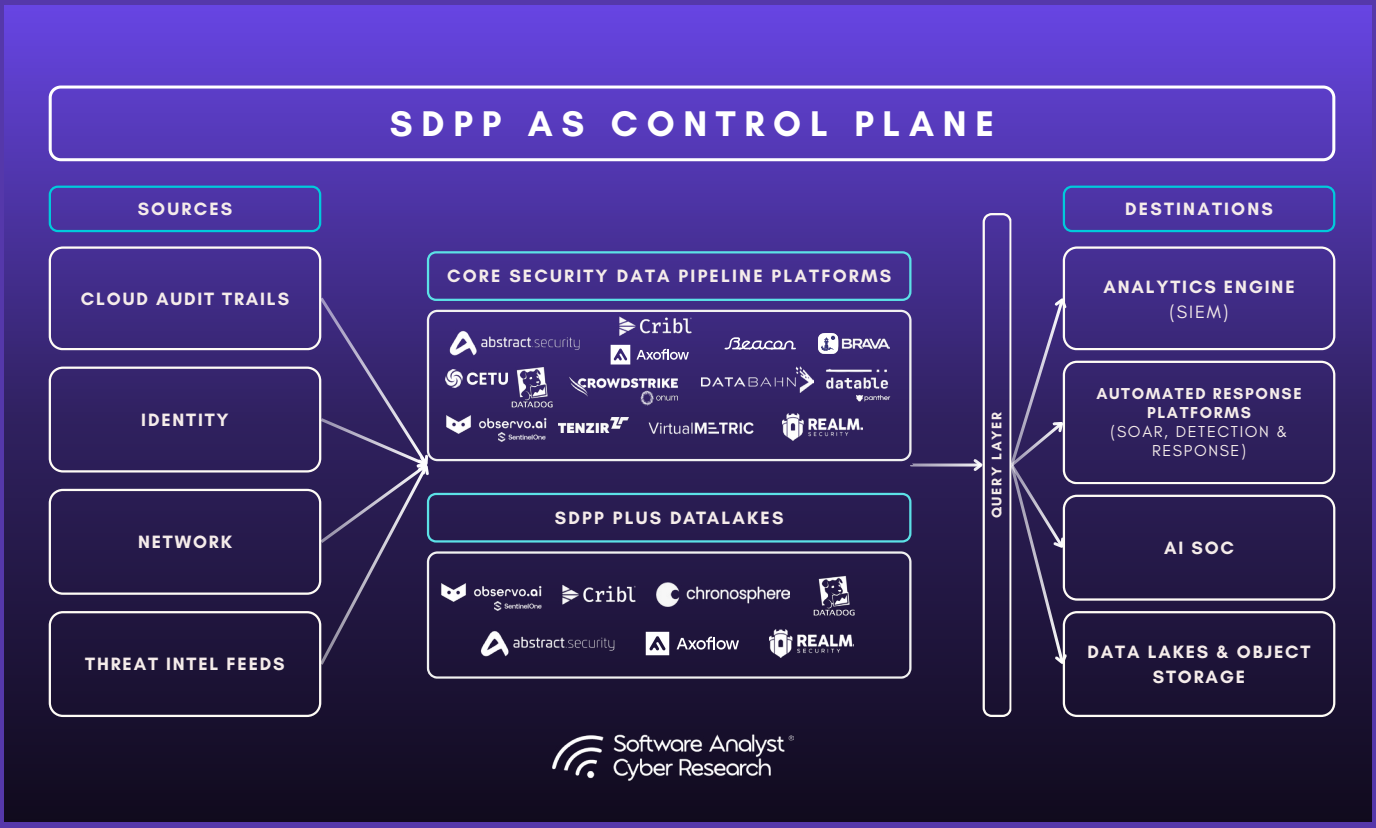
The Pipeline Becomes the Control Plane

Security data pipelines are often misunderstood as ETL for security, simple brokers that route data from point A to point B. But modern platforms deliver far more. **They are becoming the control plane** for how security teams manage, govern, and trust their telemetry across SIEM, detection, response, AI, observability and long-term analytics.

Security leaders today face rising data volumes, constant schema changes, noisy logs, silent data dropouts, inconsistent enrichments and mounting SIEM and data lake costs. Traditional ingestion or basic filtering cannot keep up. What is emerging is a new class of platforms designed specifically for security data. These platforms reduce noise without losing security context, normalize and enrich at scale, auto generate parsers, detect schema drift, monitor data source health including silent failures and apply AI to make the pipeline self optimizing.

Across the industry, what is clear is this: Security data pipeline platforms are moving from helpful optimization to the foundational control layer of the SOC architecture. They sit at the center of the architecture and shape how every downstream tool performs.

Below is a concise breakdown of the key capabilities these platforms bring and the innovations security leaders should watch as the category evolves.



Readers Note: Some of the features mentioned in the section below may only be offered by more advanced security data pipeline platforms. In the vendor section, you will find a detailed description and an in-depth evaluation that makes it easier for security leaders to compare and understand what each platform provides.

Core Pipeline Capabilities

Core pipeline capabilities across most security data pipeline platforms include the following

Advanced Data Reduction Beyond Simple Filtering

Data reduction is not just about shrinking data volume. In security, it means preserving investigative value while eliminating noise and unnecessary cost. This section covers how modern pipelines intelligently reduce data without weakening detection fidelity.

Early pipeline wins came from cost savings, but reduction has become much more intelligent than dropping fields.

What SDPPs can actually do

- Context aware suppression that removes duplicates or repetitive events while preserving indicators and security context
- Conditional reduction at both field and event level
- Adaptive sampling that dynamically adjusts sampling rates based on peak ingestion times
- Payload trimming that removes non security relevant metadata like verbose debug fields or oversized payloads that add cost but not value
- Schema aware reduction that preserves detection relevant fields while trimming high volume noise
- Summarization and metricization that convert chatty logs into compact metrics without losing investigative value
- Priority based reduction that adjusts logic by log type
- Real time shape correction that transforms data in proper formats before they hit downstream systems

Together, these techniques ensure data reduction is cost efficient and security aware rather than blind trimming.

Emerging Innovations

- Dynamic reduction tuned by threat context or incident state
- AI assisted reduction recommendations based on historical alerting patterns
- User configurable reduction tiers aligned to detection criticality
- Automated validation to ensure reductions never strip fields needed for investigations

Why it matters

Security teams reduce SIEM spend while keeping the fidelity needed for investigations. Leaders repeatedly said they want tools that help them do more with less without degrading security.

Normalization and Schema Discipline

Normalization ensures that every log source speaks a consistent language. This allows detections, analytics, and investigations to work reliably across diverse destination systems. Schema discipline prevents breakage and enables large scale correlation.

Nearly every practitioner called this the top priority.

What SDPPs can provide

- Automatic normalization into standards such as OCSF, ECS, UDM or custom schemas
- Schema drift detection when a data source silently changes formats
- Automatic parser creation using AI for new versions and undocumented logs
- Consistent field naming across all data sources to unlock SIEM content and correlation

Emerging innovations

- AI generated parsers based on sample logs and intended destination schema
- Automated detection of unexpected new fields or missing required fields
- Version aware normalization that adapts when vendor log formats update
- Normalization confidence scoring to flag risky transformations

Why it matters

If data does not show up clean and consistent, SIEM, XDR, SOAR, UEBA, AI SOC and detections all suffer. Good data unlocks the entire detection library.

Contextual and Threat Intel Enrichments

Raw logs lack the context analysts need. Enrichment adds meaning, giving logs identity, asset and threat relevance so alerts and queries become more accurate and actionable.

SDP platforms act as enrichment hubs, adding rich context in stream, to strengthen analysis at destination.

Examples

- Environmental context such as GeoIp, cloud account or region
- Identity context such as user, department and privilege level
- Asset context including owner, business app and criticality
- Threat intel matches for IPs, domains and hashes

Emerging innovations

- Pre enrichment policies that vary based on log type or threat level
- Inline lookup optimization for high speed enrichment
- Automated asset tagging based on behavioral patterns
- Dynamic enrichment paths that enrich only when detection relevance is high

Why it matters

Enrichment turns raw events into signals analysts can act on. By adding context such as identity, asset and threat context in the pipeline, teams reduce triage time, improve correlation quality and make AI driven use cases more reliable without adding extra steps later.

Intelligent Routing and Multi Tier Storage Control

Not all data should be treated equally. Intelligent routing ensures each log is sent to the right place at the right cost tier while maintaining flexibility across SIEMs, data lakes, and analytics tools.

This is where the pipeline becomes the control plane.

What SDP Platforms provide

- Route hot, warm and cold based on log value
- Split streams to multiple SIEMs, detection tools or cloud lakes
- Apply different reduction schemas by destination

Emerging innovations

- Price aware routing where users can split pipeline routes to choose storage based on cost differences across cloud providers
- Vendor agnostic SIEM migration paths

Why it matters

Routing and storage decisions directly drive cost, performance and flexibility. Intelligent routing lets security teams control where data lives, keep hot paths fast for investigation and avoid being locked into a single SIEM or storage vendor.

Intelligent Integration Health Monitoring

Noise is not the only enemy, sometimes silence is a bigger threat. Security teams need to know not just what is happening, but whether critical telemetry is flowing at all times. Monitoring for noise, errors and silent dropouts ensures visibility gaps do not turn into undetected incidents.

What SDP Platforms provide

- Detect silent quitting of sources - Integration health at source level
- Monitor ingested volume against historical baselines
- Alert on pipeline stalls or destination issues

Innovations to note

- Automatic discovery of newly active or inactive sources
- Health scoring of each data source over time
- Behavioral baselines for normal telemetry flow
- Automated response actions when a source goes dark
- Detect sudden drops in fields or event types

Why it matters

If critical sources go dark or degrade, SOC metrics may still look healthy while real blind spots grow. Source and pipeline health monitoring makes telemetry reliability visible so teams can trust their coverage claims and respond quickly to gaps.

AI Assisted Pipelines

AI brings speed and automation to pipeline tasks that were historically slow, manual and error prone. Practitioners are now growingly comfortable with the idea of AI use within pipeline platforms, The high value application of AI on pipeline platforms is not to be compared as similar to autonomous SOC claims. Instead of replacing analysts, AI in the pipeline reduces operational burden by providing pipeline recommendations, accelerates onboarding of new data sources and strengthens data quality before detections even begin.

AI directly addresses several long standing pain points. First, onboarding new log sources is too slow, especially during platform migrations or when connecting new environments. Security leaders noted that the speed at which they can ingest and normalize data determines how quickly they can address issues at destination platforms like SIEMs. AI generated parsers and automated normalization drastically shorten the time from raw logs to usable telemetry.

Second, many leaders stressed that high alert volume is often not due to weak SOC workflows but because prerequisite work in data quality, clustering and correlation is incomplete. AI in the pipeline helps reduce this noise by automatically grouping related events, generating cleaner schemas and ensuring that logs arrive enriched and structured, which lowers the alert queue burden downstream.

Third, teams repeatedly warned about silent failures in telemetry. AI powered baselining and anomaly detection on data flow can identify when sources go dark, when formats drift or when volumes shift abnormally, addressing a critical visibility gap.

Early innovations

- AI generated parsers
- AI driven pipeline creation
- Automated anomaly detection in transit
- Semantic classification of log types

Innovations to note

- Recommendations for pipeline optimization
- AI analysis that validate schema drift and transformation status
- Predictive detection of missing log integrations
- Automated rerouting when pipeline detects destination health failures

Why it matters

AI assisted pipelines absorb repetitive engineering work and constant change in vendor formats. That frees scarce security engineers and analysts to focus on detections, investigations and architecture rather than plumbing.

Unified Security Data Control Plane

As capabilities converge, security pipelines are becoming the strategic control layer that governs how telemetry is shaped, enriched and used across the entire SecOps stack.

What it provides

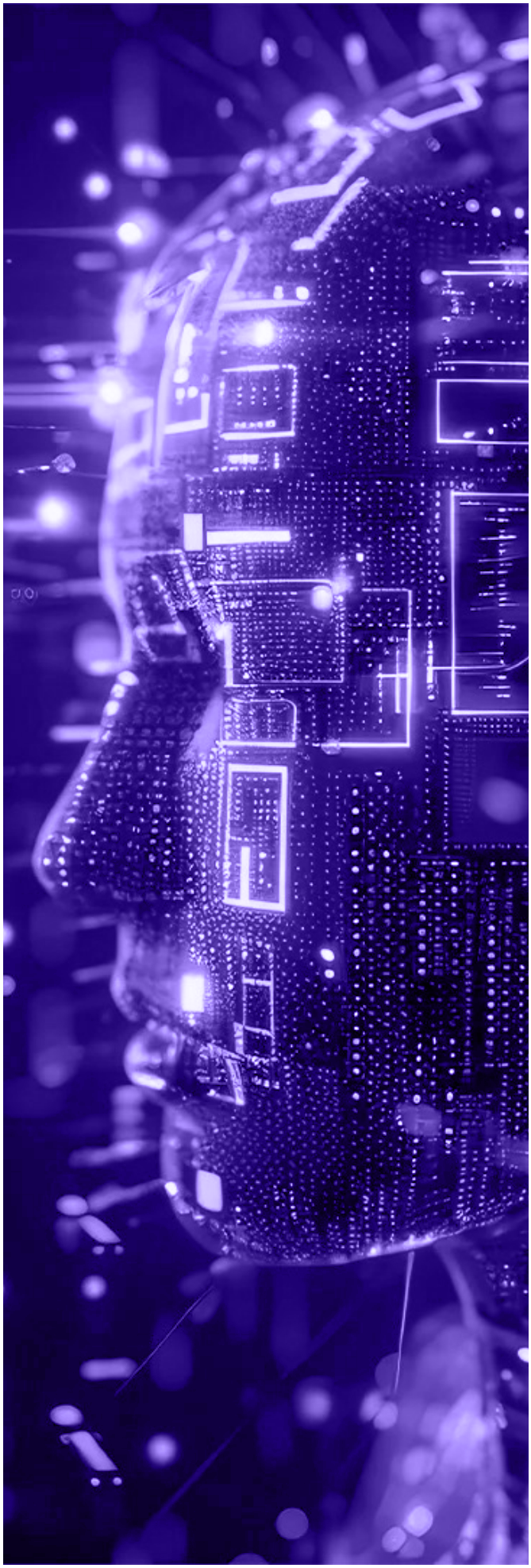
- Central governance of data quality
- One place to enforce schema, reduction and routing policy
- A foundation for consistent AI and analytics
- Control plane APIs for external orchestration
- Policy as code for data governance
- Unified dashboards showing security, cost and performance impacts
- Automated end to end lineage tracking for every event

Why it matters

Treating the pipeline as a unified control plane gives CISOs one place to govern data quality, cost and access. This foundation makes it easier to evolve tools, adopt new analytics and AI, and respond to regulatory or business changes without constantly reworking integrations.

When reduction, normalization, enrichment, schema governance, routing, data health and AI automation converge, pipelines become the central control plane for deciding how telemetry is used in end systems.

Recent acquisitions show that SIEM vendors understand this. The pipeline is the strategic chokepoint. **Whoever controls the data layer influences the entire SOC stack.**



Emerging Trends

Here are some emerging trends we see across some of these modern security data pipeline platforms –

Deployment and Distribution Flexibility

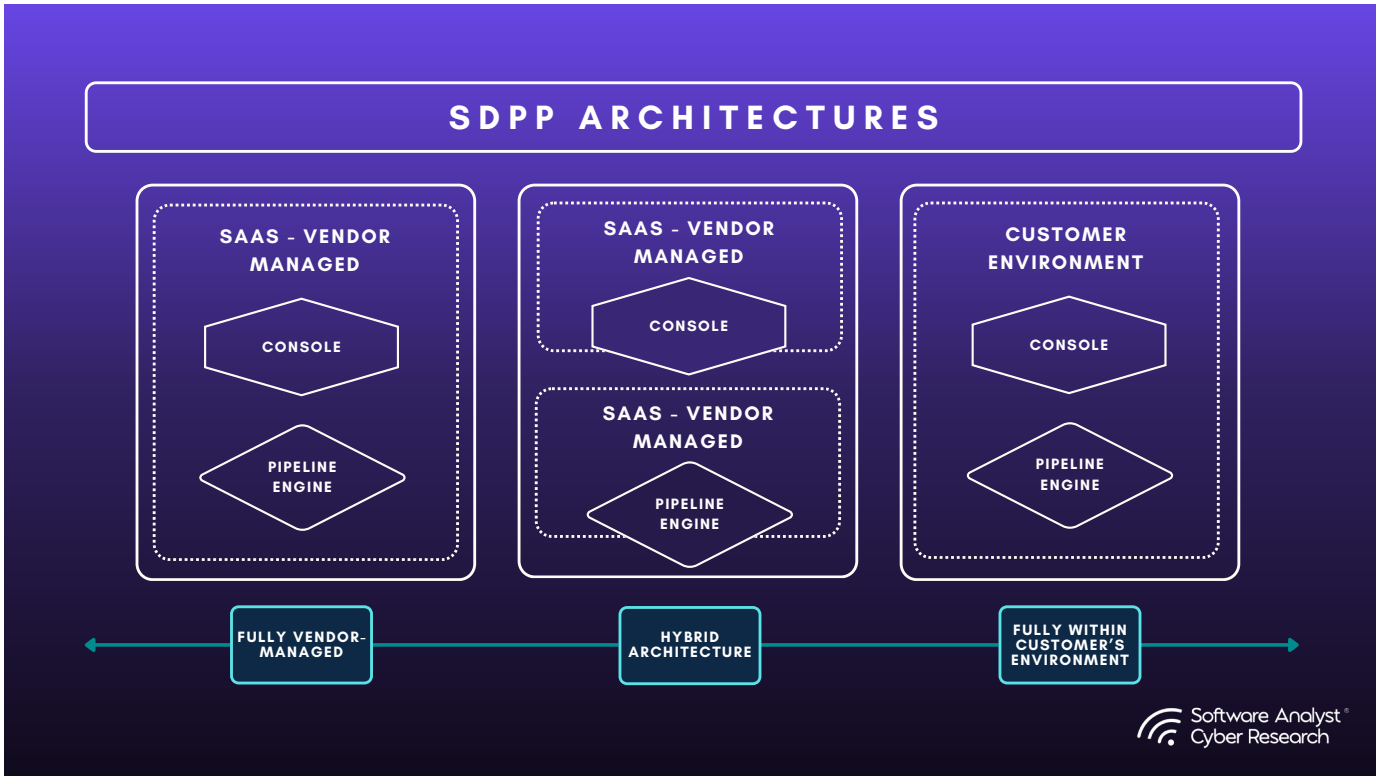
We see these platforms offering flexible deployment options, typically using a split model with a Control Plane and a Pipeline Engine. Many of these vendors also offer multi-tenancy to support MSSPs and large enterprises.

Advanced Normalization, Enrichment, and Context Fabric

Normalization and enrichment are becoming richer and more automated across vendors. Abstract normalizes into multiple schemas, enriches with identity, asset, vulnerability, and threat intel, and auto-corrects drift with ASE. Databahn transforms data into CIM, OCSF, UDM, ASIM, LEEF, and more using AI while enriching with STIX/TAXII threat intel. Axoflow auto-classifies logs, applies schema mapping, and enriches with metadata. Beacon aligns to ECS, OCSF, CIM, and UDM while combining cross-source context through Recipes. CeTu provides AI-assisted normalization with lookup enrichment and threat intel overlays. Brava enriches telemetry with attack-simulation context, relevancy scoring, and MITRE mappings. Cribl enriches data through lookups, Redis, GeoIP, and DNS, with schema drift detection forthcoming.

Intelligent Routing and Multi-Destination Control

Routing decisions are becoming value-based and policy-aware across vendors. Abstract recommends routing based on detection value and cost. Axoflow uses classification labels to drive automated routing. Beacon's AI-guided posture directs logs to SIEM, data lakes, or cold storage based on importance. CeTu's Zoe assistant selects routes tied to analytic relevance, cost, and detection needs. Databahn's Cruz AI evaluates query patterns and detection impact to recommend tiering and routing paths. Brava routes high-efficacy logs



forward while summarizing or filtering low-value data. Cribl Stream provides granular routing from any source to any destination, using Copilot to generate logic from plain language.

Integration Health and Coverage Insights

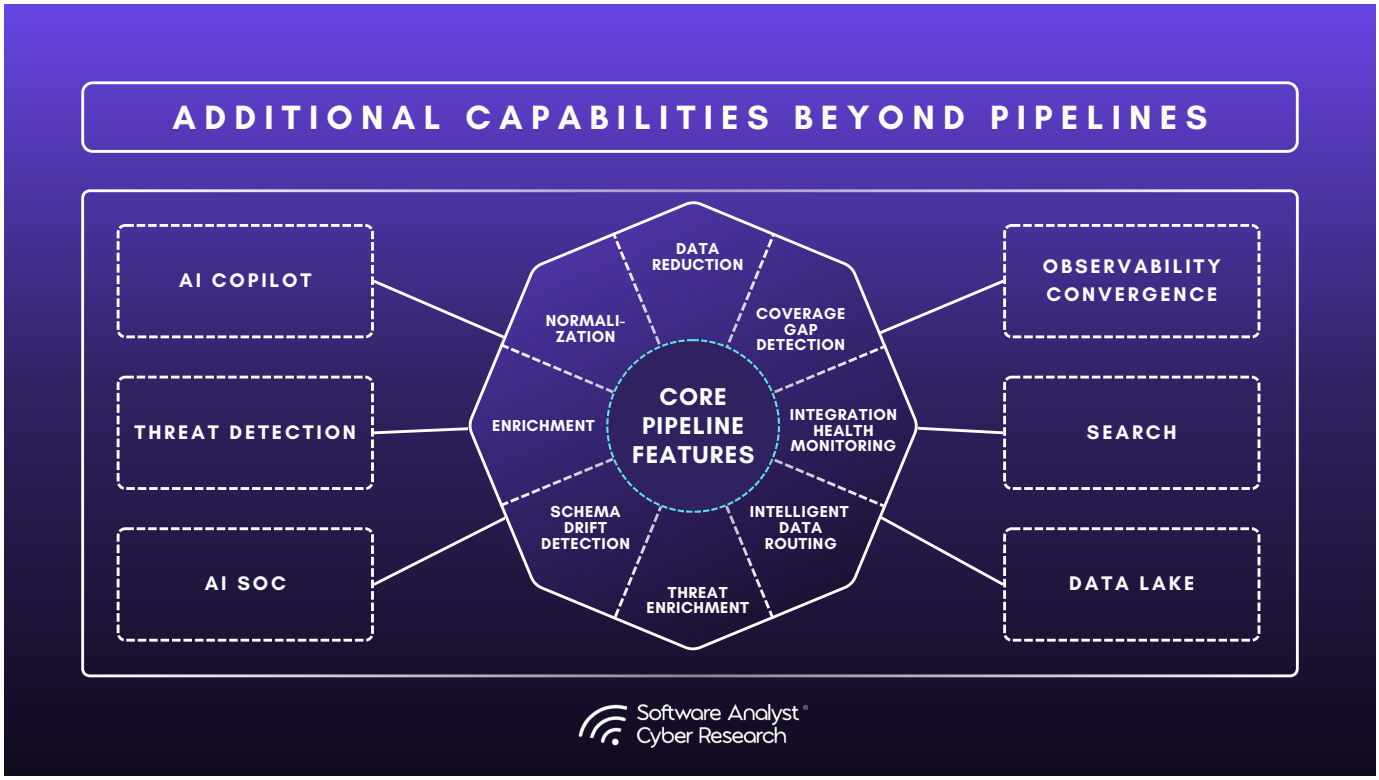
Vendors now offer deep insight into data coverage, stability, and silent failures. Abstract detects silent dropout, schema drift, and volume anomalies with automated parser correction. Beacon's Logging Posture highlights missing telemetry and coverage gaps using its Collectopedia knowledge base. Databahn scores source health based on quality, completeness, drift, and destination stability. Axoflow alerts on missing sources, unexpected new sources, and message drops. Brava maps coverage gaps using attack simulation aligned to MITRE techniques. CeTu VISION analyzes SIEM coverage and highlights blind spots across environments. Cribl Insights surfaces backpressure, drops, latency, and health issues across Stream, Edge, and Lake deployments.

AI-Assisted Pipeline Management

AI is maturing into a core operational layer in nearly all vendors. Databahn emerged as a leader in their AI capabilities and maturity. Abstract's ASE generates parsers, manages drift, builds pipelines, and enriches detections. Axoflow uses supervised AI for classification, schema mapping, and natural-language pipeline creation. Databahn's Cruz automates parser generation, correction, routing intelligence, and ecosystem-specific model transformations. Cribl's Copilot assists with schema mapping, routing logic, and query generation. Beacon applies agentic reasoning to Recipes, posture, schema mapping, and normalization. CeTu's Zoe and DEPTH engines power routing decisions, drift detection, and pattern intelligence. Brava uses AI to evaluate telemetry efficacy through attack simulations and relevancy scoring tied to detection strength.

SDP PLUS Platforms

In addition to core pipeline features, we are seeing an emerging trend where pure-play SDP platforms are envisioned to expand beyond traditional pipeline capabilities. These include in-house data lake option with tiered storage, AI assisted capabilities, threat detections and analytics in the pipeline layer, federated searches and querying across SIEMs and Datalakes, Observability convergence and AI SOC like capabilities. Here are some of the features we saw among the vendors we analyzed.



Data Lakes and Tiered Storage

Vendors increasingly offer storage and replay layers that extend pipelines into long-term retention. Abstract provides Lake Villa with hot, warm, and cold tiers and real-time querying. Cribl offers Cribl Lake and Lakehouse for open-format retention and fast access to recent data. Axoflow includes AxoStore, AxoLocker, and AxoLake as part of its multi-layer storage design. Databahn offers optional tiered lake storage for customers needing centralized history. Brava supports seamless retrieval from low-cost storage directly through the SIEM. CeTu does not mandate a lake but provides a unified architecture that can route to object stores or archival platforms. Beacon avoids owning storage but enables routing to cold tiers across customer-controlled buckets.

Search, Querying, and Federated Visibility

Search and query capabilities are expanding directly from the pipeline layer. Cribl Search enables search-in-place across S3, Edge, Lake, and external object stores. CeTu offers cross-system querying that works across SIEMs and data lakes without a query language. Brava embeds natural language querying within the SIEM and retrieves cold data transparently. Databahn supports micro-indexing for rapid search across raw pipelines. Abstract allows real-time queries over normalized data through Lake Villa. Axoflow's debugging and inspection tools show raw and parsed data side-by-side to support source comprehension. Beacon provides transformation previews and exploratory data analysis to validate pipeline accuracy.

Shifting Detections to the Stream

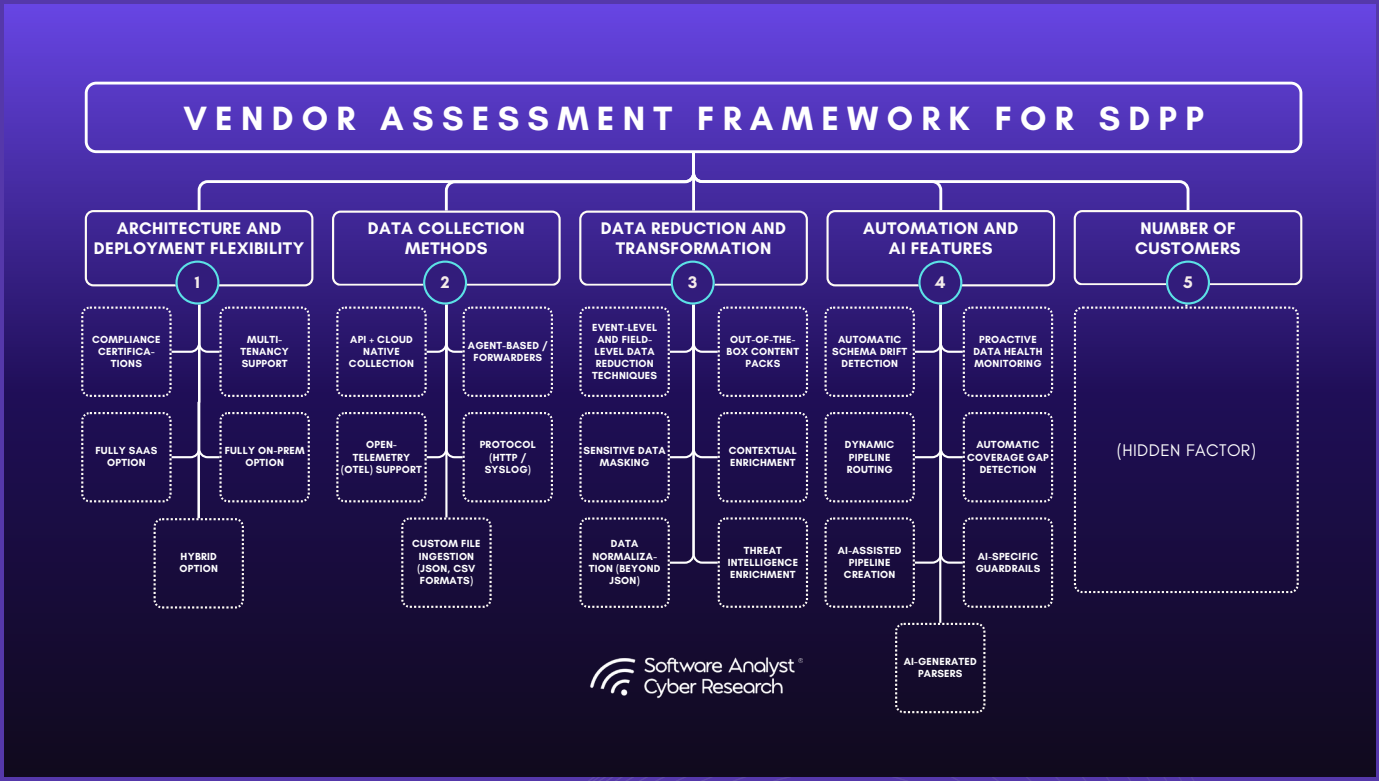
Threat detection in stream represents a shift toward identifying malicious activity as data flows through the pipeline rather than after it lands in a SIEM or data lake. The idea is to evaluate events in real time, applying lightweight correlation, IOC checks, and contextual signals before logs are indexed, which can reduce dwell time and provide earlier visibility into suspicious behavior. This approach also allows detections to carry enriched context downstream, improving the quality of alerts and investigations. While teams appreciate the speed and proximity to the source, most see in-stream detection not as a replacement for SIEM analytics but as a complementary layer that helps surface high-value signals, reduce noise, and begin the investigative process earlier. Vendors such as Abstract Security, VirtualMetric, acquired pipeline platforms and Tenzir already offer some threat detection capabilities. Realm Security is another entrant that plans to add this to its roadmap in the near future.

Evaluation Framework

Security data pipeline platforms deliver lower SIEM and storage costs at the minimum, but they also provide higher detection quality, better data governance, faster investigations, safer AI adoption, more resilient telemetry and freedom from vendor lock-in.

Most importantly, they shift teams from reactive ingestion problems to proactive control of the entire data lifecycle. They are not a sidecar tool. They are becoming the backbone of modern SecOps architecture.

In order to evaluate these vendors in depth, we conducted several deep-dive platform demos, used detailed questionnaires with linked evidence and screenshots to validate responses, and interviewed their customers to confirm our findings. Here are the broader categories under which the vendors were evaluated:



Vendors

Disclaimer: The image above is not depicted by exact ranking. Please see the ranking below. Spreadsheet with Technical and GTM assessment. Note that in-depth details are not contained within the sheet to maintain clean format but the details are covered within in each vendor's section to prove how these rankings were made.

From our in-depth analysis, we found the pipeline platforms to rank as below -

Overall Market and Category Leader: Cribl

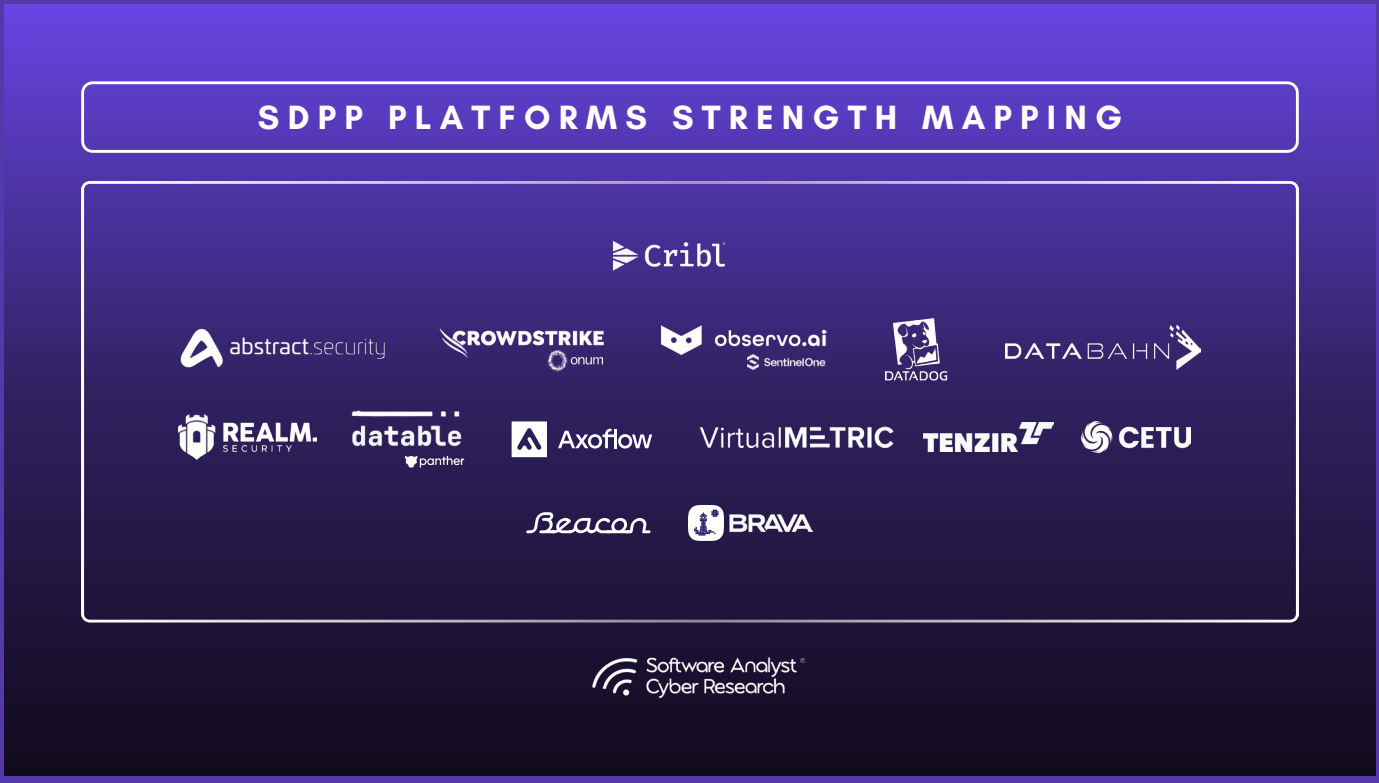
Pipeline Leaders: Databahn, Datadog OP, Abstract Security, Observo AI, Onum

Emerging Leaders: Cetu, VirtualMetric, Tenzir, Axoflow, Datable, Realm Security

Innovators: Brava, Beacon Security

The following top vendors were evaluated with a thorough platform demo, in-depth questionnaire (the answers of which were verified via demo and screenshots) and direct customer / practitioner feedback.

In alphabetical order and no particular ranking, all details on vendors below —



REALM.
SECURITY

Realm.Security

Realm is a promising newcomer to the security data pipeline platforms industry. They raised $15M in Series A just recently on Oct 8 2025. Realm positions itself as a tightly integrated, multi-tenant cloud control plane built around single-tenant data pipelines, persistent queues, and ML or LLM-driven relevance and normalization.

Voice of the Customer

We were able to interact with a customer of Realm to understand their use cases and experience with Realm. Here is what they said –

Life before Realm

"We needed a vendor agnostic platform to allow us to have control over the data that we send to our SIEM platform. We needed this to make sure that we could evaluate other SIEM platforms and quickly/easily transition to those platforms once a new one was selected. We needed to migrate/duplicate some of our data for long term retention purposes. We also have critical log sources that contain PHI/PII that we were unable to ingest into our SIEM as our SIEM is hosted in a third party cloud platform (defacto HIPAA violation) - Realm allows us to clean those logs of PHI/PII so that they can be ingested and we do get the alerting we need. We also have a small SIEM team, so the AI capabilities within Realm to intelligently trim log data to save on ingestion costs is hugely beneficial. Short term it will enable us to drive down SIEM costs, migrate SIEM platforms and keep PII/PHI out of our SIEM. Longer term the SIEM is likely going to be less of a focus for the industry as AI bots will be able to do most of that work, so Realm will still play a pivotal role in serving as the pipeline of the data that would once go explicitly to a SIEM to instead move to alternative data lakes, while making sure that those logs are continuing to be trimmed and maintained."

Most used capabilities within Realm

"Data migration, data duplication, data scrubbing and intelligent log cleanup."

What they'd like to see more of

"Going beyond pipeline management and a SIEM focus and potentially adding functionality where it can send AI bots into our various sources and extracting data without needing to explicitly send logs to a SIEM."

Architecture and Deployment Maturity

Realm offers flexible deployment options. The platform separates its multi-tenant control plane (Console) from single-tenant data pipelines (Pipeline Engine). The control plane handles configuration, policy, and observability, while each customer's dedicated pipeline processes and stores data within its own secure tenancy.

Marketplace: Hyperscaler partnerships; marketplace listings planned.

MSSPs: Private, undisclosed partnerships.

Compliance: SOC 2 certified

Pricing

Daily Ingestion. Consumption-based model

Pricing Assistance

- Pricing calculator: Realm provides a pricing calculator for pricing assistance.

Data Collection and Integrations

1. OTEL Collectors:

Realm supports OpenTelemetry (OTEL) collectors for on-premises environments. These collectors standardize telemetry before forwarding to Realm pipelines.

2. Agents and Forwarders:

Lightweight agents and forwarders handle continuous data streaming from hosts, applications, and network sources.

3. API-Based (Agentless) Integration:

Realm enables agentless data ingestion through REST and streaming APIs. This method suits SaaS and cloud-native tools where direct API connectivity is preferred over deploying local agents.

4. Cloud Syslog Streams:

Cloud syslog integration allows ingestion from cloud-based security and infrastructure services that expose syslog endpoints.

5. Push Integrations:

Realm supports both push and pull data models for cloud services. Sources can push telemetry directly to Realm endpoints.

Number of integrations: 15 out of the box, with expanded possibilities because of collection methods.

Core Pipeline Capabilities

Capability	Details
1. Data Reduction Capabilities	• Data Reduction and Intelligent Filtering • Aggregation • Deduplication • Event Discards and • Field Filtering
2. Out-of-the-Box Content Packs	While not “content packs”, each customer receives rule recommendations within the data fabric for every integration configured. These recommendations serve as the default mechanism for managing data transformations, providing out-of-the-box guidance without additional setup.
3. Data Normalization	Realm currently normalizes data to JSON and plans to add support for OCSF, ECS, and CIM in the first half of next year.
4. Enrichment	Security context and business context **Threat Intel Enrichment**: Yes. Used for log reduction and filtering rules.
5. Schema Drift Detection	Realm detects schema drift and automates required parser updates for customers via support.
6. Threat Detection Content	Currently under development.
7. Intelligent Data Routing	Realm automates routing recommendations based on data type and destination use.

Additional Pipeline Capabilities

Intelligent Log Detection: Automatically identifies log types and routes them to the correct pipelines.

Secure Filtering and Testing: Implements a Dev-Test-Run framework with event capture to visualize transformation impact.

Persistent Queuing: Ensures no data loss during downstream outages and resumes delivery when systems recover.

Pipeline Building Experience

Drag and Drop interface: Provides drag-and-drop source and destination objects for faster pipeline configuration.

AI Maturity

Realm applies ML and LLM models across ingestion, filtering, normalization, and configuration to reduce reliance on expert scripting and support cost-aware routing. The platform's intelligence layer incorporates inference mapping, combining statistical machine learning, security context, and incident response best practices to improve data quality and minimize alert fatigue before forwarding to SIEM or XDR systems.

Realm's knowledge graph is used in production to generate personalized transformation rule recommendations, independent of its developing agentic workflows. The intelligence layer is designed to feed future agentic capabilities to support autonomous cross-source correlation.

Realm can autonomously parse datasets ingested into the platform, and the team is currently evaluating whether to expose this configuration for direct customer interaction

Data Privacy with AI

Tenant level isolation based on architecture.

Identity and Access Management

Includes SSO, SAML, custom OIDC, MFA, and granular RBAC controls.

Integration Health Monitoring

Capability	Description
Integration Health Monitoring	Provides real-time integration active / error health status with alerts that identify likely causes and offer remediation recommendations to accelerate resolution.
Coverage Gap Analysis	Not available.
Schema Drift Detection	Automatic detection of schema drift. Parser updating to fix the schema would require manual intervention which is also supported by Realm's support team to do so on behalf of the customers.

Additional Capabilities beyond Pipelines

Data Lake

Realm provides a data storage capability known as Data Haven, designed as a raw archival layer with retrieval and rehydration functionality. The system was built to support machine querying from the outset, enabling efficient access to historical data for analytics and AI-driven workflows. Realm positions Data Haven as both a cost-optimized extension of SIEM storage and a future-ready layer for agentic or AI SOC integrations, allowing customers to retain more data within budget while preparing for emerging automation and AI use cases.

Vision

Realm frames its direction around affordable ingest expansion, AI-assisted pipeline configuration, and a haven-plus-replay (data lake) architecture that will eventually support AI SOC enablement. Leadership is direct about the roadmap being guided by customer needs and transparent about areas of the product still maturing.

Analyst Take

Here's what we see as major strengths and opportunities for improvement for Realm –

Strengths:

In conversations with leadership, their goal is clear on where they want to go and where they don't. Realm's vision aligns closely with what buyers care about from Pipelines platforms, most: cost, speed, and reliability. Their focus on reducing ingest spend and simplifying deployment addresses a clear market pain point. The use of ML and LLMs for filtering and normalization addresses an area where many teams still struggle to automate and is in the direction where leading pipeline platforms are heading with AI capabilities. The combination of a multi-tenant control plane with single-tenant pipelines, supported by year-long pipeline reporting, gives early teams the maintenance flexibility and centralized visibility in distributed organizations they need as they scale.

Areas to Watch:

Realm remains in the early stage of its buildout, with a smaller integration footprint and will need time to reach broad source coverage. Marketplace listings and MSSP partnerships are still developing, which may create short-term friction for cloud-based procurement. Realm will need to accelerate supported normalization formats and invest in advance AI capabilities such as automatic coverage gap analysis and parser updates. Advancement will hinge on how quickly Realm expands integrations and delivers marketplace availability, while maturing AI and pipeline feature set. Practitioners should watch near-term progress on these fronts.

Software Analyst®
Cyber Research

business personal

softwareanalyst.io

SACR®

Trusted research. Sharp insights. Real conversation.

CISO	VENDOR
SECURITY TEAMS	INVESTORS