Publication date:
04 Jul 2025
Author(s):
Andrew Braunberg, Principal Analyst

Market Landscape: Cybersecurity Data Fabrics 2025

Omdia
by informa techtarget

Table of Contents:

Summary

It has long been said that cybersecurity is a data management problem; however, until recently, enterprises have been hesitant to invest in a strategic reexamination of their security data architectures. Data fabrics have emerged as a viable path forward for holistic data democratization and cost containment initiatives.

Catalyst

In this Market Landscape, Omdia explores the use of data fabrics to create comprehensive frameworks for managing cybersecurity data. Data fabrics can be delivered as standalone solutions that security teams can build DIY analytic solutions on top of, or delivered as the foundations of broader platforms that support one or more security use cases directly.

The approach has proven useful in a diverse set of applications, including customer service in retail and banking, patient care in healthcare, and supply chain management in manufacturing. Building data fabrics to manage security data is a newer phenomenon, but security use cases are multiplying quickly. Omdia sees great utility in the application of data fabrics and data pipeline management tools (often key components in actualizing fabrics) to augment existing security controls, particularly security information and event management (SIEM), and proactive security platforms such as exposure management.

Omdia view

Collecting and managing security data has proven to be a difficult and persistent problem. Most security operations centers (SOCs) are built around mature threat detection, investigation, and response (TDIR) solutions, commonly centered on next-generation SIEM. These solutions are workhorses for log ingestion and analysis; however, they are expensive to operate, particularly with respect to costs associated with data retention, and they often do not provide the flexibility in data use and federation needed by modern enterprises.

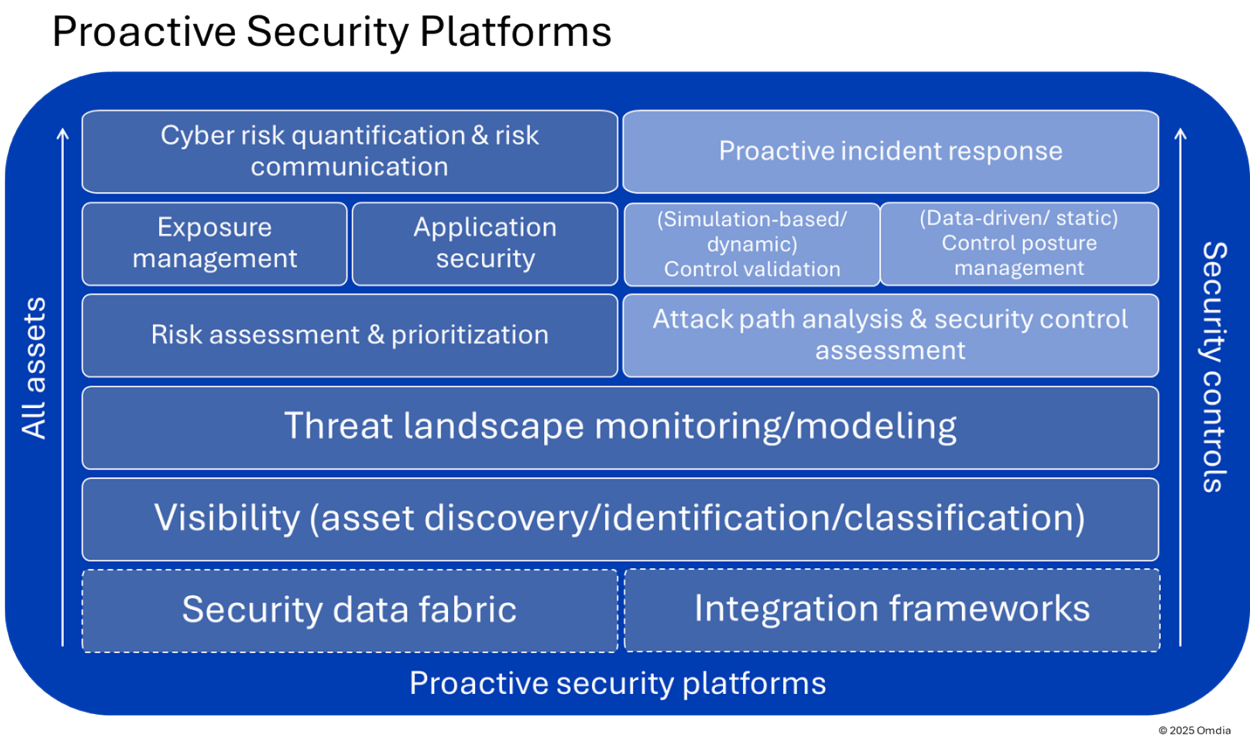
Data pipeline management tools have emerged, with SIEM augmentation as a primary use case. The purpose of these tools is to decouple the frontend data discovery and processing tasks from the SIEM architecture. By moving more intelligence into the pipeline, data storage decisions can be made before data routing decisions.

For organizations to embrace a fully federated data architecture, however, takes a more strategic rethinking of how security data is discovered, processed, connected, and consumed. This is where data fabrics can play an important role in modernizing how organizations utilize security data. Data fabrics typically support a broad set of capabilities in addition to pipeline management, most notably data governance features such as data lineage.

The overall market for security data management solutions continues to attract new entrants, investors, and acquisition interest. Omdia is tracking dozens of vendors that position themselves as solution providers of data fabric, data pipeline management, or both. Data fabric vendors have been

particularly focused on proactive security use cases, and data fabrics can act as a foundation for a host of Continuous Threat Exposure Management applications (see **Figure 1**).

Figure 1: Data fabrics are a foundational component of a new class of proactive security platforms



Source: Omdia

Key messages

- Security platforms are widely touted as a solution to the data silo problem in cybersecurity. Platformization is not, however, the only solution to the difficulty in fully utilizing data across a heterogeneous collection of security controls. Data fabrics provide a framework for abstracting the complexity of security data management and enabling the delivery of the right data to the right analytic engine at the right time.
- Data fabrics are in part a reaction to the perceived shortcomings of previous attempts (e.g., SIEMS) to consolidate security data. Data fabrics leverage automation and open schemas to simplify management of security data and democratize the use of that data by enabling a "many-to-many" flow of data from any source to any destination.
- At a minimum, data fabrics should support data discovery, inventory, orchestration, and governance. Data pipeline management is a key element of broader data fabric frameworks. And a key benefit of data pipeline tools is the ability to manage data sources and destinations independently of specific proprietary platforms, processes, and methods.
- Data fabric frameworks have been used successfully in many non-security use cases. While the market for security data fabrics is more nascent, numerous startups have emerged in the last several years that are productizing data fabrics for security data.

Recommendations

Recommendations for enterprises

- Data fabrics should be attractive to organizations for security use cases if they have siloed and fragmented data, require a diverse set of real-time analytics applications, and are moving toward a broad strategy for security automation powered by artificial intelligence (AI).
- Key benefits of the adoption of data fabrics include democratization of security data through a more open and flexible data architecture. Enterprises should examine the estimated cost containment benefits of decoupling security data from proprietary architectures and enabling use-case-specific routing and storage of data.
- The term data fabric remains a bit fuzzy, and enterprises are advised to perform appropriate due diligence on potential solution providers to ensure support for specific use cases. Prospects should have a full understanding of the current and road-mapped support for AI-powered automation features.

Recommendations for technology vendors

- Vendors that have seen the most success selling these technologies have taken one of two approaches. Either they have gone to market with a limited set of use cases that support a detailed argument about cost containment and data storage options, or they have positioned as a data platform that supports additional storage (e.g., cyber asset attack surface management [CAASM]) or analytic applications (e.g., exposure management).
- The CAASM category has been slow to grow, in no small part because vendors often had difficulty demonstrating value from inventory capabilities alone. CAASM tools were early to market and could not predict the rise of AI and broad contextual analysis, but can be viewed as early efforts to produce data fabrics for security use cases.
- Vendors should view customer concerns related to security data management as an opportunity. The total cost of ownership of security data has become a serious concern that needs to be addressed. Vendor lock-in, as well as an inability to find low cost options for storing lower value data, have become major drivers of customer dissatisfaction. Legacy security vendors have ignored these concerns long enough for a host of startups to emerge.

Market status

Definition and characteristics

A data fabric is designed to reduce the complexities associated with data movement, transformation, integration, and analysis. It operates across platforms, data processing methods, data delivery methods, and data storage architectural approaches, and often relies on a deep understanding of the metadata associated with each data source. Data fabrics should address metadata discovery, analysis, contextualization, and relationship mapping.

Data pipeline management, data observability, and data fabrics

Omdia views data pipeline management as a subsegment of data fabrics. Data pipeline management supports tools and processes for moving data. Data observability is a related but separate discipline that focuses on monitoring and analyzing the quality and reliability of data and data pipelines. Standalone data observability solutions are outside the scope of this report.

By decoupling data sources from data destinations and adding the ability to process and operate on the data before it reaches a destination, data pipeline management tools solve several pressing challenges for security teams. Perhaps most importantly, these tools help security teams more effectively and efficiently manage the large and growing volume of security telemetry that is needed for TDIR. As noted, a primary use case for these tools is augmenting legacy SIEM to better manage the data storage costs associated with those solutions.

Data pipeline management vendors include Cribl, DataBahn, Observo AI, Tarsal, and Tenzir. Not surprisingly, SIEM and extended detection and response (XDR) vendors are also working to simplify data pipeline tasks. For example, Splunk has developed two tools to help customers build data pipelines. These are the Edge Processor and the Ingest Processor. And CrowdStrike has partnered with Cribl to create CrowdStream, a cloud-hosted version of Cribl Stream, which is available through CrowdStrike Falcon LogScale.

Data fabrics typically include data pipeline management features, and there is considerable overlap between vendors positioned as data pipeline management and data fabric solution providers. However, data fabrics deliver a broader set of capabilities to address a richer set of use cases. Omdia believes that a data fabric solution should provide the following baseline capabilities:

Data catalog

- Data cataloging is the process of organizing or inventorying data assets, and a data catalog allows organizations to create a comprehensive mapping of their data holdings.
- A key benefit of a data catalog is its data search and discovery features.

Metadata management

- Allows exploration of networks and relationships to discover connections between data.
- A machine learning-based discovery engine can scan and discover data assets across the enterprise.
- A metadata knowledge graph is a structured representation of information that uses entities and relationships to organize and connect data.

Security data orchestration

- **Data ingestion**
  - **Batch processing:** Scheduled processing of large volumes of data.
  - **Real-time streaming:** Continuous processing of data in support of real-time analysis and response.
  - **Application programming interface (API) integrations:** Use of APIs to enable real-time data exchange between different systems or applications.
- **Normalize, parse, and transform:** These traditional data processing tasks are increasingly automated with AI tools to normalize, parse, and transform security data into a common

data format such as open cybersecurity schema framework (OCSF), common event format, common information model, or to a security data lake or other data storage solutions.

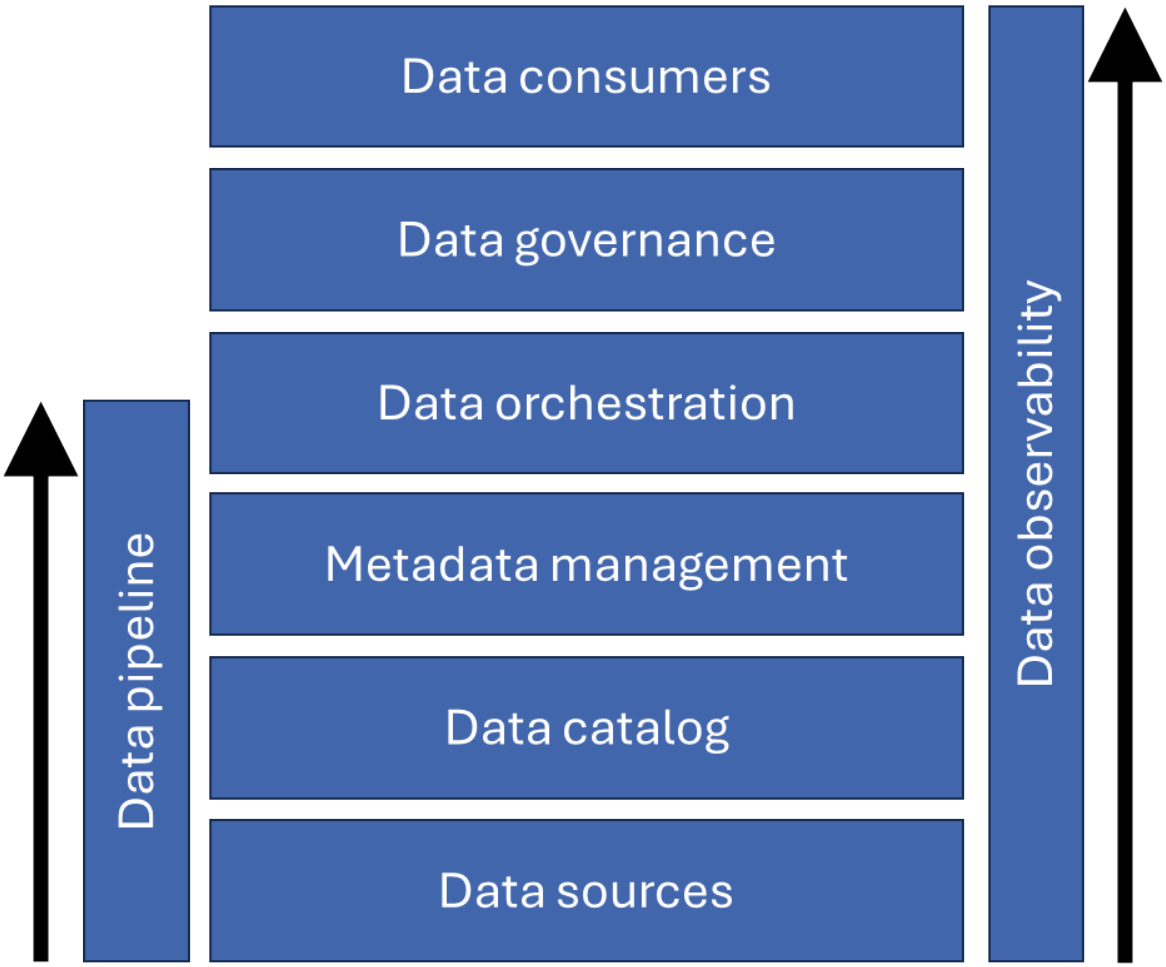
- **Data enrichment:** This traditional security data processing task is also increasingly automated with AI tools. Enrichment is critical to fully understanding and leveraging the value of any dataset. An example is correlating asset data with the latest threat data to determine specific common vulnerabilities and exposures or Indicators of Compromise that impact those assets.
- **Data forking:** Determine the appropriate use case(s) for data before routing it to storage. For example, forking could allow logs determined to have threat detection value to be routed to an analytic engine, while other logs could be sent directly to cold storage for compliance purposes.

Data governance

- **Data lineage:** Monitor the provenance of data by tracking where it comes from and how it moves through a network to provide full traceability as well as accountability. Ongoing data monitoring to ensure the quality, reliability, and performance of all data.
- **Data security features:** Access control and role-based permissions, for example.
- **Data quality:** Monitor and alert on data schema drift and track the consistency, completeness, and relevance of the security data.
- **Compliance:** Ensuring compliance with regulatory mandates and data sovereignty requirements.

Data fabrics typically include an app or web-based frontend to support the configuration of data sources and systems and the creation of data models. The actual work is done with a backend engine that automatically keeps track of each data source connection, storage options, and processing requirements. An important benefit of data fabrics is simplifying the data management expertise needed from security teams. For example, data fabrics should abstract the details of create, read, update, and delete transactions from the end user.

Figure 2: Security data fabric architecture



Source: Omdia

Market dynamics

Key capabilities and vendor landscape

Deploying data fabrics should be viewed as a strategic initiative that can leverage the entire security stack. By taking a metadata-focused approach to data management, data fabrics simplify and often automate data discovery, access, and governance. In fact, this holistic metadata-level view enables a consistent governance strategy across data sources and unifies data access and data security policies through centralized policy management. Finally, and most importantly, data is made available for real-time analytics by any approved analytic engine or application. This democratization of data usage can significantly increase the value of existing security telemetry and reduce vendor lock-in and reliance on proprietary data schema.

It is perhaps ironic that SIEM augmentation is a primary use case of these data management tools, because the SIEM was one of the first security products designed to solve the data silo problem by bringing together log data to TDIR activity in the SOC. Consolidating data into proprietary SIEM solutions, however, has created its own set of constraints, particularly with respect to managing the costs associated with data storage.

Beyond log data

The event-focused log data that SIEMs process is but a subset of the mix of structured, semi-structured, and unstructured data that can be managed through a more comprehensive data fabric. This additional data could include vulnerability, identity, and asset datasets in support of compliance analytics, security analysis, and metrics tracking. Data fabrics often take an asset-based (e.g., devices, users) approach to organizing data and understanding data relationships, but data does not need to physically move into a centralized data store. A data virtualization tool can be used to create a virtual data layer that integrates metadata from each data source and avoids physically moving the data. This can be a significant benefit, particularly if internal politics make it difficult to create a centralized data repository across departments.

The use of graph databases can simplify data virtualization. Graph database technology is already well established within a broad set of security solutions. Examples include: CrowdStrike Threat Graph, Cisco SecureX, Microsoft Security Exposure Management, Palo Alto Networks Cortex XDR, SentinelOne Purple AI, and Trend Micro Vision One. There are also a host of standalone graph database solutions, including Neo4j, TigerGraph, and Amazon Neptune. Graph databases are particularly good at visualizing and analyzing the connections between data in real time, in support of threat detection, attack path analysis, and risk prioritization.

Future directions

The process for ingesting data into a cybersecurity fabric involves direct connections and integrations with a broad range of cybersecurity, infrastructure, development, and application tooling. This process often requires custom connector development and a reasonably high level of ongoing maintenance, as APIs for access change over time.

The future of data ingestion for cybersecurity data fabrics will focus on additional automation, with AI-powered agents playing a key role. They will learn how to connect, collect, and analyze the available data within each target for use within the context created by the fabric. They will possess the intelligence required to proactively discover all relevant data sources in a digital ecosystem, including previously unknown assets. The agent will be able to understand data semantics, adapt to evolving or changing APIs, and establish and learn new connections dynamically, resulting in a decrease in maintenance and total overhead of operations. Already today, AI agents are being used very successfully to automate data normalization tasks, such as standardizing data in OCSF.

Data governance will also benefit from additional automation. This will include automated data lineage tracking, data access controls, and data quality monitoring. A general trend will be to continue to automate and shift data management tasks left by moving data processing tasks close to data sources.

As pipelines get smarter, it may also be possible to include detection logic in pipeline processing. Pipelines could begin to function as real-time threat sensors that execute detection rules (e.g., Sigma) on data before they are pushed to a SIEM or another analytic engine.

Market outlook

Optimizing security data management strategies has become a top-of-mind issue for many security leaders. The expected benefits include cost containment, improved security posture, and increased flexibility in data utilization.

The democratization of security data requires a shift away from proprietary data architectures and the enablement of a data management strategy that promotes many-to-many data collection, routing, and usage. Both security data fabrics and data pipeline management solutions work to achieve a level of data democratization.

Data pipeline management vendors have built a market primarily around SIEM augmentation; however, the core features and functions in these tools are important components in more comprehensive data fabric solutions, which would typically include additional data governance and data security features. Omdia observes that many data fabric vendors are building suites of analytic applications on top of their core fabric platforms, often in support of proactive security use cases. This trend highlights the uncertainty in how buyers will consume data fabrics in the long term, either as standalone, independent solutions or as foundations of larger, comprehensive analytic platforms.

Vendors to watch

Abstract Security

Abstract Security, founded in 2023, has built a solution to better centralize the management of data operations and security analytics. Abstract's platform supports the routing, transformation, and enrichment of security telemetry to any number of sources (e.g., cloud storage, SIEMs, and data platforms). The solution supports optimized cost controls through automated routing of data to hot, warm, and cold storage tiers. Additionally, the platform allows organizations to identify threats with advanced analytics and correlation tools.

Comcast/Databee

Databee was founded in 2022 to bring a commercial version of Comcast Cybersecurity's security data fabric platform to market. The data fabric solution is designed to address security, risk, and compliance use cases in Fortune 500 organizations. Supported functionality includes: Continuous Compliance and Risk Management for automated controls reporting across, for example, PCI, NIST, and DORA ((Digital Operational Resilience Act); AI-powered asset/owner discovery and inventory to uncover blind spots in configuration management databases and automate discovery of applications and devices; Vulnerability and Asset Exposure Management to correlate vulnerability scanner results with asset context; Threat Detection and Response for automated OCSF normalization and entity resolution; and cost-optimized data storage.

Cribl

Cribl was founded in 2018 by three ex-Splunk employees. The company was an early proponent of the need for SIEM augmentation through data pipeline management. Pipelines are at the core of Cribl's Stream processing that enables a logical sequence of functions to process data events and perform various operations. Functions are pieces of JavaScript code that Cribl Stream performs on each event that passes through a pipeline. In 2023, CrowdStrike introduced CrowdStream, which is based on Cribl technology.

Deepwatch/Dassana

Deepwatch, a managed detection and response vendor, acquired data fabric vendor Dassana in February 2025. The Deepwatch managed security platform provides threat intelligence and management capabilities with a dynamic risk-scoring engine to correlate alerts and provide contextualized responses to threats. By integrating Dassana's technology into its platform, Deepwatch assists customers in the shift from a centralized SIEM model to a distributed data architecture. Deepwatch's Open Security Data Architecture leverages AI, security data, intelligence, and human expertise to reduce risk through threat detection and remediation.

Mondoo

Mondoo was founded in 2020. The company's core IP is a security data fabric, which provides the foundation for the open source "policy as code engine," cnspec. Mondoo has built a unified security platform that leverages cnspec to identify, prioritize, and address risks across on-prem, cloud, software as a service (SaaS), and endpoints. The platform supports a host of use cases, including vulnerability management, unified security posture management, and compliance automation. The platform also supports policy as code, or the practice of codifying policies into machine-readable formats. Security, compliance, and cost control policies can be expressed as code and applied across digital environments, including cloud, on-prem, Kubernetes, SaaS, endpoints, and software development lifecycle environments.

Onyxia Cyber

Onyxia Cyber was founded in 2022, and its flagship product is the AI-Powered Preemptive Cyber Defense and Planning Platform. The company has a suite of solutions on top of the platform in support of data-driven program management, exposure management, predictive and actionable AI program insights, security stack management, and ROI and cost optimization. The platform is built on a security data fabric that leverages AI and machine learning to continuously analyze resources to connect data that would otherwise remain siloed. The goal is to create universal, real-time, and historical data that can be consumed in any environment to support strategic and operational security initiatives.

Tenzir

Tenzir's goal is to ensure that an organization's security data is a strategic asset, not a vendor-controlled burden. The company was founded in 2017 and is headquartered in Hamburg, Germany. The Tenzir platform is built on an open source pipeline engine for security teams. Data flows can be built into pipelines that allow security teams to build, manage, and automate security data workflows with modular, composable tools. Pipelines can be built to collect, shape, normalize, optimize, enrich, store, replay, and route any security telemetry data.

Zscaler/Avalor

Zscaler acquired Avalor, and its security data fabric solution, in March 2024. Avalor's Data Fabric for Security automates data ingestion, normalization, and enrichment to provide security teams with a comprehensive and up-to-date view of their cybersecurity posture. Zscaler is building a suite of analytic applications on top of the security data fabric (e.g., Asset Exposure Management, Risk360, Unified Vulnerability Management), and the solution will also accelerate Zscaler's implementation of AI into its portfolio.

Appendix

Methodology

Omdia performed numerous vendor interviews to understand the current market for data fabrics and data pipeline management solutions. Omdia used revenue estimates from its SecOps market tracker to confirm claims of market traction.

Author

Andrew Braunberg, Principal Analyst, SecOps

askanalyst@omdia.com

Citation policy

Request external citation and usage of Omdia research and data via citations@omdia.com.

Omdia consulting

We hope that this analysis will help you make informed and imaginative business decisions. If you have further requirements, Omdia's consulting team may be able to help you. For more information about Omdia's consulting capabilities, please contact us directly at consulting@omdia.com.

Copyright notice and disclaimer

CONTACT US

omdia.com

askananalyst@omdia.com