GIGAOM

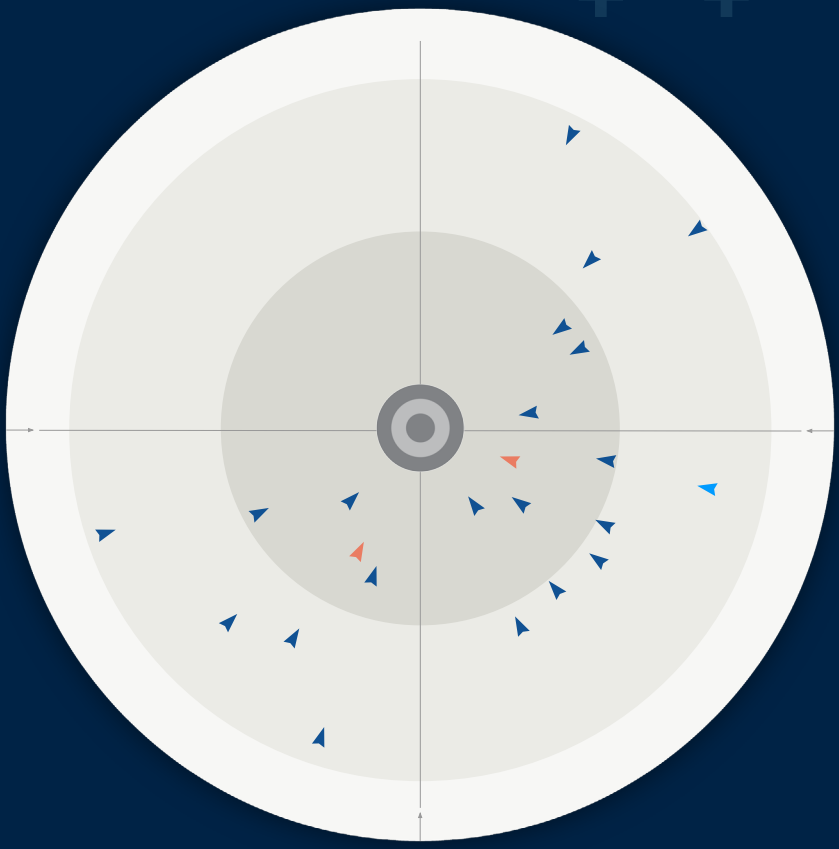
In partnership with:

Hitachi Vantara

Unstructured Data Management v6

James Brown

CLOUD, INFRASTRUCTURE & MANAGEMENT




GIGA
OM

Unstructured Data Management v6

James Brown

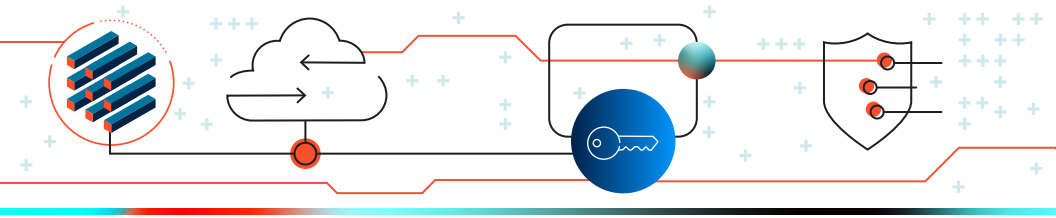
Table of Contents

01 Executive Summary

This GigaOm Radar report is commissioned by Hitachi Vantara

THE GENERATIVE AI REVOLUTION has transformed unstructured data from a passive storage burden into an enterprise's most valuable and riskiest asset. This data, including emails, documents, videos, chat logs, and sensor feeds, now comprises over 90% of all information, holding the key to competitive advantage, operational efficiency, and breakthrough innovation. However, without a modern management strategy, it also represents a significant source of security breaches, compliance failures, and uncontrolled costs. A robust unstructured data management (UDM) platform is no longer optional; it is the foundational control plane for AI-powered enterprises.

Prospective buyers should assess UDM solutions based on the key criteria that define a modern, intelligent platform with seamlessly integrated capabilities. A leading solution must begin with AI-powered discovery and classification to autonomously identify and contextualize data at petabyte scale, automatically surfacing sensitive information and business-critical content. This core intelligence then feeds deep analytics and insight extraction, leveraging large language models (LLMs) and computer vision to unlock the "who, what, when, and why" within the data. This valuable insight is protected by a zero trust security posture, which implements granular access controls and continuous monitoring to ensure data is secure by default. Complementing this access control and monitoring is automated governance and compliance, which enforces data lifecycle and privacy policies as code. All these capabilities must operate seamlessly across a unified hybrid and multicloud data fabric, delivering a single pane of glass for managing data, policies, and security wherever data resides.



Business Imperative

For C-suite executives, the business case is clear and compelling. A leading UDM strategy is a direct investment in the organization's future, driving tangible results across the board. This manifests first as proactive risk reduction, achieved by automating sensitive data protection and continuously monitoring for compliance gaps and security threats. It also unlocks significant cost optimization by eliminating redundant, obsolete, and trivial (ROT) data and automating storage tiering. Critically, in today's landscape, it accelerates revenue and innovation by making high-quality, business-relevant data securely accessible to fuel AI/ML models and analytics teams. Finally, these automated systems enhance overall operational efficiency through self-service data access and a drastic reduction in the manual labor required for data handling and forensics.

This is our sixth year evaluating the unstructured data management space in the context of our Radar reports. This report builds on our previous analysis and considers how the market has evolved over the last year.

This GigaOm Radar report examines 23 of the top unstructured data management solutions and compares offerings against capabilities (table stakes, key features, and emerging features) and nonfunctional requirements (business criteria). It provides an overview of the market, identifies leading unstructured data management offerings, and helps decision-makers evaluate these solutions so they can make a more informed investment decision.

A robust unstructured data management (UDM) platform is no longer optional; it is the foundational control plane for AI-powered enterprises.

02 Market Categories and Deployment Types

TO HELP PROSPECTIVE CUSTOMERS find the best fit for their use case and business requirements, we assess how well UDM solutions are designed to serve specific target markets and deployment models (**Table 1**).

For this report, we recognize the following market segments:

- **Infrastructure-led:** This segment focuses on high-performance, resilient storage architectures managed by IT infrastructure teams. Buyers prioritize durability, scalability, and cyber recovery. Purchase decisions center on total cost of ownership, hardware compatibility, and operational simplicity.
- **Platform-centric:** Solutions offer unified data management and security under one control plane. Buyers seek integration breadth, governance, and automation. Purchase considerations include ecosystem fit, platform maturity, and long-term roadmap stability.
- **Intelligence-led:** This segment centers on metadata analysis, automation, and AI-driven insight. Buyers prioritize contextual data visibility, risk reduction, and compliance automation. Purchase decisions emphasize accuracy, analytics depth, and interoperability with existing security and governance tools.

In addition, we recognize the following deployment models:

- **Customer-managed:** Deployed in the customer's own infrastructure for maximum control and security. Ideal for regulated or latency-sensitive environments but requires internal management of updates, scalability, and maintenance.
- **SaaS:** Delivered as a vendor-managed service with automatic updates, scalability, and predictable costs, reducing IT overhead and accelerating deployment.
- **Hybrid cloud:** Combines local control with cloud-based intelligence and scalability. Enables centralized policy management and analytics while keeping data local for performance, cost, or compliance needs.

Table 1 components are evaluated in a binary yes/no manner and do not factor into a vendor's designation as a Leader, Challenger, or Entrant on the Radar chart (**Figure 1**).

"Target market" reflects which use cases each solution is recommended for, not simply whether that group can use it. For example, if an SMB could use a solution but doing so would be cost-prohibitive, that solution would be rated "no" for SMBs.

Table 1. Vendor Positioning: Target Market and Deployment Model

PLATFORM	TARGET MARKET			DEPLOYMENT MODEL		
	INFRASTRUCTURE-LED	PLATFORM-CENTRIC	INTELLIGENCE-LED	CUSTOMER-MANAGED	SAAS	HYBRID CLOUD
Aparavi	-	-	✓	✓	✓	-
Arcitecta	-	✓	-	✓	✓	✓
BigID	✓	✓	✓	✓	✓	✓
Cohesity	-	✓	-	✓	✓	✓
CTERA	-	✓	✓	✓	✓	✓
Data Dynamics	-	-	✓	✓	✓	-
Dell Technologies	✓	✓	-	✓	-	✓
Druva	-	✓	-	-	✓	-
Hammerspace	-	✓	✓	✓	-	-
Hitachi Vantara	✓	✓	-	✓	✓	✓
HPE	✓	✓	-	✓	-	✓
IBM	✓	✓	-	✓	✓	✓
Komprise	-	-	✓	✓	✓	✓
Nasuni	-	✓	✓	-	-	✓
NetApp	✓	✓	-	✓	-	✓
Panzura	-	✓	✓	-	-	✓
Pure Storage	✓	-	-	✓	-	✓
Quantum	✓	-	-	✓	-	-
Qumulo	✓	-	-	✓	-	✓
Rubrik	-	✓	✓	✓	✓	✓
Scality	✓	-	-	✓	-	-
StrongLink	-	✓	✓	✓	-	-
Varonis	-	-	✓	✓	✓	-

Source: GigaOm 2026

03 Decision Criteria Comparison

ALL SOLUTIONS INCLUDED in this Radar report meet the following table stakes—capabilities widely adopted and well implemented in the sector:

- Data and metadata collection
- Storage connectivity
- Unified search and retrieval
- Policy-based data movement
- Basic data classification
- Reporting
- APIs and extensibility

Tables 2, 3, and 4 summarize how each vendor in this research performs in the areas we consider differentiating and critical in this sector. The objective is to give the reader a snapshot of the technical capabilities of available solutions, define the perimeter of the relevant market space, and gauge the potential impact on the business.

- Key features differentiate solutions, highlighting the primary criteria to be considered when evaluating a UDM solution
- Emerging features show how well each vendor implements capabilities that are not yet mainstream but are expected to become more widespread and compelling within the next 12 to 18 months
- Business criteria provide insight into the nonfunctional requirements that factor into a purchase decision and determine a solution's impact on an organization

These decision criteria are summarized on the next page.

Key Features:

- **AI-driven data intelligence:** This feature uses AI to autonomously understand the business context of data, moving far beyond basic metadata search. This is critical for unlocking the hidden value in file content and making data more valuable for business users and other systems.
- **Generative AI data curation:** This feature actively prepares an organization's data to serve as a trustworthy source for generative AI (GenAI) applications like retrieval-augmented generation (RAG). This is essential for ensuring AI models are fueled by high-quality, relevant data, which prevents inaccurate or "hallucinated" results and maximizes the business value of AI investments.
- **Data security posture management (DSPM):** DSPM provides continuous, proactive visibility into where sensitive data is located, who has access to it, and how it's being used. In an era of persistent cyber threats and stringent privacy regulations, this proactive approach to security is critical for preventing costly data breaches before they happen.
- **Automated policy enforcement:** This feature moves beyond passive reporting to actively and automatically enforce data governance policies for lifecycle, retention, and security. Automation is the only viable way to manage data at a petabyte scale while ensuring consistent compliance, reducing operational costs, and freeing up skilled IT staff for more strategic initiatives.
- **Cyber recovery orchestration:** This feature helps an organization detect, respond to, and orchestrate a rapid and clean recovery from a cyberattack like ransomware. Given that attacks are now inevitable for most large organizations, the speed, accuracy, and reliability of the recovery process has become a critical business continuity function.
- **Enterprise workflow integration:** This feature measures a platform's ability to integrate deeply into the broader enterprise IT and security ecosystem, sharing intelligence and automating actions across different systems. This transforms the UDM platform from a data silo into a central nervous system for data-centric operations, increasing the value of existing technology investments.
- **Customizable AI/ML classification:** This allows an organization to train the UDM platform to recognize its unique, business-specific, or proprietary data types. This is a crucial differentiator for moving beyond generic personally identifiable information (PII) detection to accurately classifying an organization's most valuable intellectual property, such as engineering plans, legal documents, or scientific research.

Table 2. Key Features Comparison

KEY FEATURES	AVERAGE SCORE	AI-DRIVEN DATA INTELLIGENCE	GENERATIVE AI DATA CURATION	DATA SECURITY POSTURE MGMT (DSPM)	AUTOMATED POLICY ENFORCEMENT	CYBER RECOVERY ORCHESTRATION	ENTERPRISE WORKFLOW INTEGRATION	CUSTOMIZABLE AI/ML CLASSIFICATION
Aparavi	2.1	***	-	**	****	*	**	***
Arcitecta	3.4	****	****	**	*****	*****	*****	****
BigID	4.1	*****	****	*****	****	**	****	*****
Cohesity	3.7	***	****	****	*****	*****	****	***
CTERA	3.4	***	***	***	***	****	***	**
Data Dynamics	2.0	****	***	****	****	***	***	***
Dell Technologies	2.0	***	****	***	****	*****	****	**
Druva	2.1	****	***	***	****	*****	***	**
Hammerspace	3.1	**	***	*	*****	****	****	**
Hitachi Vantara	3.0	****	*****	****	****	****	****	****
HPE	4.1	**	***	**	****	*****	****	**
IBM	3.6	***	****	****	*****	*****	****	****
Komprise	2.3	****	****	**	*****	*	*****	***
Nasuni	4.1	***	***	***	***	*****	***	**
NetApp	4.0	***	***	****	****	****	****	***
Panzura	3.6	*	-	*	***	*****	***	*
Pure Storage	3.4	*	***	*	***	****	***	*

Key features differentiate solutions, highlighting the primary criteria to be considered when evaluating a UDM solution.

Quantum	**3.1**	*	*	*	****	***	***	*
Qumulo	**3.1**	*	**	*	***	***	***	*
Rubrik	**4.1**	**	**	*****	*****	*****	****	***
Scality	**3.1**	*	**	*	****	***	***	*
StrongLink	**2.0**	***	***	*	*****	***	*****	**
Varonis	**3.7**	****	**	*****	****	**	*****	****

***** Exceptional ****Superior *** Capable ** Limited * Poor – Not applicable or absent

Source: GigaOm 2026

Emerging Features

Conversational data intelligence: This feature allows users to interrogate and manage their data estate using a natural language, conversational interface. This is important because it democratizes data intelligence, empowering nontechnical users to get complex questions answered without needing to learn a specialized query language.

AI data provenance and lineage: This feature provides the ability to track the complete origin (provenance) and journey (lineage) of data used to train AI models. This is critical for ensuring the trustworthiness of AI, providing an audit trail for regulatory compliance, and protecting models from data poisoning.

Unified data fabric extension: This feature extends the UDM platform's visibility beyond unstructured files to correlate that content with structured data from enterprise databases and applications. Breaking down this final silo provides a truly holistic view of enterprise information, unlocking unprecedented business insight.

Table 3. Emerging Features Comparison

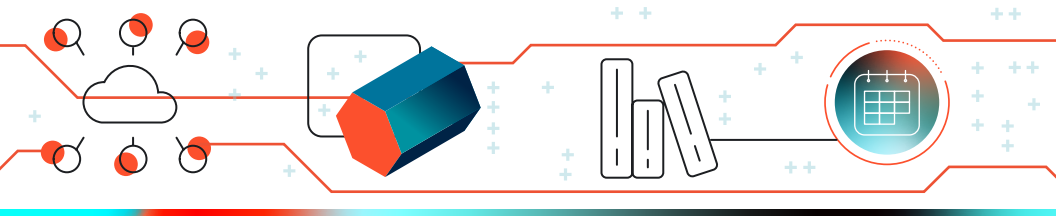
EMERGING FEATURES	AVERAGE SCORE	CONVERSATIONAL DATA INTELLIGENCE	AI DATA PROVENANCE AND LINEAGE	UNIFIED DATA FABRIC EXTENSION
Aparavi	0.0	–	–	–
Arcitecta	2.3	–	****	***
BigID	4.3	*****	***	****
Cohesity	2.3	****	*	**
CTERA	1.3	***	–	*
Data Dynamics	1.7	–	***	**
Dell Technologies	2.7	***	*	****
Druva	2.0	***	*	**
Hammerspace	1.3	–	***	*
Hitachi Vantara	2.7	–	***	*****
HPE	3.0	–	****	*****
IBM	4.0	***	****	*****
Komprise	1.7	–	***	**
Nasuni	0.7	–	*	*
NetApp	1.3	–	**	**
Panzura	0.7	–	*	*
Pure Storage	0.7	–	*	*
Quantum	0.7	–	*	*
Qumulo	0.7	–	*	*
Rubrik	2.0	***	*	**
Scality	0.7	***	*	*
StrongLink	2.0	*	****	**
Varonis	1.7	*	**	***

***** Exceptional **** Superior *** Capable ** Limited * Poor – Not applicable or absent

Source: GigaOm 2026

Business Criteria

- **Cost transparency:** This criterion assesses the platform's ability to provide clear visibility into the costs of unstructured data and to provide the tools to actively optimize that spend. This is critical because uncontrolled data growth is a major driver of escalating storage and cloud costs.
- **Scalability:** This metric evaluates the platform's ability to handle growth in data volumes, file counts, and throughput without a corresponding degradation in performance. It measures the architectural capacity to support the enterprise's exponential data expansion smoothly and cost-effectively.
- **Flexibility:** This criterion assesses the platform's ability to support diverse data types, storage systems, cloud environments, and deployment models. It measures the solution's capacity to fit into a heterogeneous enterprise ecosystem without imposing rigid constraints or vendor lock-in.
- **Ease of use:** This criterion measures the ease of deploying, managing, and using the platform at scale, which directly impacts operational efficiency and user adoption. A simple platform reduces the need for specialized skills, minimizes the risk of human error, and accelerates the organization's response to business and security requests.
- **Auditability:** This metric assesses the platform's ability to provide a comprehensive and immutable record of data activity to support security forensics and prove regulatory compliance. This is a nonnegotiable requirement for any enterprise operating in today's landscape of stringent privacy laws and persistent security threats.
- **Extensibility:** This criterion evaluates how well the UDM platform integrates into and extends the value of an organization's existing technology investments through its APIs and integrations. This is critical for preventing the solution from becoming a data silo and instead transforming it into a central intelligence layer for the entire enterprise.



Emerging features show how well each vendor implements capabilities that are not yet mainstream but are expected to become more widespread and compelling within the next 12 to 18 months.

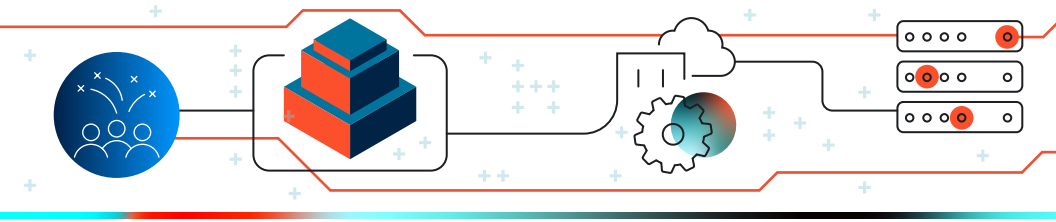
Table 4. Business Criteria Comparison

BUSINESS CRITERIA COMPARISON	AVERAGE SCORE	COST TRANSPARENCY	SCALABILITY	FLEXIBILITY	EASE OF USE	AUDITABILITY	EXTENSIBILITY
Aparavi	4.0	****	****	*****	*****	****	**
Arcitecta	4.3	*****	****	****	*****	****	****
BigID	3.8	***	*****	****	***	****	****
Cohesity	4.0	****	***	****	****	*****	****
CTERA	3.7	***	****	***	****	****	****
Data Dynamics	3.2	***	***	****	***	***	***
Dell Technologies	3.0	***	***	****	**	***	***
Druva	4.2	****	*****	***	*****	*****	***
Hammerspace	4.2	****	*****	*****	***	****	****
Hitachi Vantara	4.0	*****	****	*****	***	****	***
HPE	3.3	***	****	****	***	***	***
IBM	4.0	**	*****	*****	**	*****	*****
Komprise	3.8	*****	***	***	****	***	*****
Nasuni	3.3	****	*****	**	****	**	***
NetApp	3.8	****	***	****	****	****	****
Panzura	2.8	**	*****	***	***	**	**
Pure Storage	3.3	***	***	**	*****	***	****
Quantum	3.0	***	***	***	***	***	***
Qumulo	3.0	***	***	***	****	***	**
Rubrik	4.2	****	***	***	*****	*****	*****

Scality	**3.0**	***	****	***	**	***	***
StrongLink	**4.0**	***	*****	*****	***	***	*****
Varonis	**2.5**	**	*	***	*	*****	***

***** Exceptional ****Superior *** Capable ** Limited * Poor – Not applicable or absent

Source: GigaOm 2026



Business criteria provide insight into the nonfunctional requirements that factor into a purchase decision and determine a solution's impact on an organization.

04 GigaOm Radar

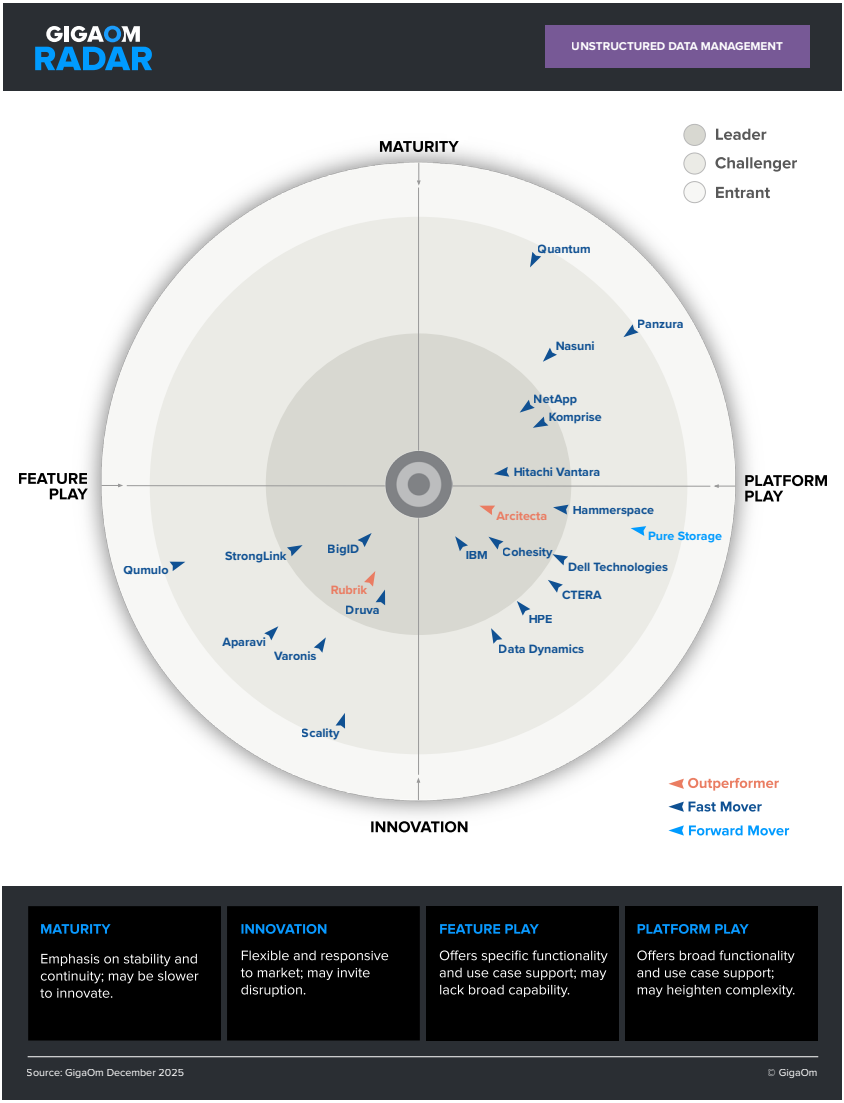


Figure 1. GigaOm Radar for Unstructured Data Management

THE GIGAOM RADAR PLOTS VENDOR SOLUTIONS across a series of concentric rings with those set closer to the center judged to be the most complete solutions. The chart characterizes each vendor on two axes—balancing Maturity versus Innovation and Feature Play versus Platform Play—while providing an arrowhead that projects each solution's expected evolution over the coming 12 to 18 months.

As you can see in **Figure 1**, the UDM market is continuing its steady evolution from storage infrastructure toward intelligence, automation, and integrated platform functionality. The chart also reveals that this transition remains uneven. While several vendors have shifted positions upward through deeper AI integration and data-driven innovation, others remain anchored in traditional, infrastructure-centric models. This year's Radar, therefore, depicts a market in flux rather than one of full convergence, with a notable shift toward Platform Plays while maintaining a visible clustering of vendors still emphasizing file and object storage reliability.

Several key observations emerge. First, more vendors now sit on the Platform Play side than on the Feature Play side. This reflects a market in which ecosystem breadth and unified architectures are becoming the primary drivers for enterprise adoption. Many Platform Play solutions are converging data protection, DSPM, and compliance automation into cohesive control planes. These organizations compete less on standalone innovation and more on integrated policy management. As enterprise buyers prioritize consolidated visibility and lifecycle automation, this segment will define the foundation for scalable AI-ready data fabrics over the next two years.

Conversely, the Feature Play side is comparatively sparse but strategic. Vendors in this hemisphere continue to address targeted problems, such as deep security governance (DSPM) or specialized workflow orchestration, without always integrating into broader storage layers. These products remain important to enterprises that are optimizing existing environments, but they highlight the fragmented nature of the market's maturity curve.

Along the vertical axis, the market displays a balanced distribution between Innovation and Maturity. While the Innovation hemisphere remains home to agile vendors emphasizing active development and "AI-first" architectures, the Maturity hemisphere hosts established leaders emphasizing operational resilience and proven scale. This year's distribution demonstrates that leadership is not confined to one side; rather, the vendors with the greatest momentum are those strategically investing in intelligence and automation, regardless of their architectural legacy.

The chart's Outperformer count remains intentionally low, reflecting how sustained over-performance in an increasingly mature market is rare. Meanwhile, Fast Movers signal measured but visible progress, bridging from infrastructure strength to intelligence infusion without disruptive transformation. Such movement underlies a market that is stabilizing, not stagnating: innovation is now systematic rather than explosive.

Year over year, the Radar captures steady consolidation rather than expansion. A few new intelligence-led vendors reinforce that AI alignment is the prevailing direction, yet most long-tenured vendors have progressed incrementally toward platform cohesion. Those that

invest early in DSPM, AI-driven analytics, and policy orchestration gain a competitive advantage, while slower adopters remain at a disadvantage.

In reviewing solutions, it's important to keep in mind that there are no universal "best" or "worst" offerings; every solution has aspects that might make it a better or worse fit for specific customer requirements. Prospective customers should consider their current and future needs when comparing solutions and vendor roadmaps.

INSIDE THE GIGAOM RADAR

To create the GigaOm Radar graphic, key features, emerging features, and business criteria are scored and weighted. Key features and business criteria receive the highest weighting and have the most impact on vendor positioning on the Radar graphic. Emerging features receive a lower weighting and have a lower impact on vendor positioning on the Radar graphic. The resulting chart is a forward-looking perspective on all the vendors in this report, based on their products' technical capabilities and roadmaps.

Note that the Radar is technology-focused, and business considerations such as vendor market share, customer share, spend, recency or longevity in the market, and so on are not considered in our evaluations. As such, these factors do not impact scoring and positioning on the Radar graphic.

For more information, please visit our Methodology.

05 Solution Insights

Aparavi: Aparavi Data Suite*

SOLUTION OVERVIEW

Aparavi is an intelligence-led software vendor focused on unstructured data discovery, classification, and lifecycle management. The Aparavi Data Suite is a single, storage-agnostic solution that scans, indexes, and analyzes data across hybrid and multicloud environments. Its core components provide data classification, a federated search index, and a policy engine for data movement and governance. This provides a focused approach centered on data intelligence, cost optimization, and compliance rather than a broad, all-in-one infrastructure solution.

Aparavi is positioned as a Challenger and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Aparavi scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The solution provides a robust policy engine that automates data lifecycle, archival, and governance actions based on metadata and classification, ensuring consistent policy execution across storage environments.
- **Customizable AI/ML classification:** It enables organizations to extend built-in classifiers and refine detection patterns for organization-specific data types, supporting more accurate and adaptable classification.
- **AI-driven data intelligence:** Aparavi uses ML analytics to correlate metadata and content attributes, improving visibility and enabling more effective decisions for compliance and optimization.

OPPORTUNITIES

Aparavi has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** The solution is focused on identifying data, not orchestrating recovery. It lacks native immutable snapshots, vaulting, or automated restore workflows.
- **Data security posture management (DSPM):** While proficient at discovering sensitive data, it falls short of a full DSPM solution, lacking advanced capabilities for mapping data exposure, analyzing access paths, and providing guided risk remediation.

- **Enterprise workflow integration:** Beyond a basic API, the platform shows little evidence of a rich ecosystem of prebuilt, bidirectional integrations with major SOAR, SIEM, or ITSM tools, which limits its extensibility.

PURCHASE CONSIDERATIONS

Licensing for the Aparavi Data Suite is typically subscription-based, often priced per terabyte of data managed or analyzed. End users buy it for specific sets of features, such as classification and policy-based tiering. It is licensed to be used in combination with other solutions (like primary storage or backup) for a best-of-breed deployment. The platform offers deployment flexibility, with options for both a fully vendor-hosted SaaS solution or a customer-managed deployment. Professional services are recommended to help configure data sources, define complex classification rules, and build initial data management policies.

USE CASES

The Aparavi Data Suite is primarily adopted to support specific, targeted use cases rather than to achieve a complete infrastructure overhaul. Key purchase drivers include analytics-driven cost optimization (identifying and tiering/archiving ROT data), data governance (enforcing lifecycle policies), and compliance (finding and managing sensitive PII and protected health information data across heterogeneous storage).

Arcitecta: Mediaflux

SOLUTION OVERVIEW

Arcitecta provides Mediaflux, a high-performance, platform-centric data fabric. Its solution is a single, integrated platform built on its proprietary XODB multimodel database, which unifies file and object data across heterogeneous storage (including disk, tape, and cloud) into a single, orchestrated namespace. This architecture provides a general approach designed to address complex, exabyte-scale data management problems in data-intensive industries. Arcitecta delivers an aggressive roadmap, exemplified by its rapid integration of native AI capabilities, and values rapid advancement and frequent updates.

Arcitecta is positioned as a Leader and Outperformer in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Arcitecta scored well on a number of decision criteria, including:

- **Automated policy enforcement:** Mediaflux delivers a powerful orchestration engine capable of managing complex, multistage workflows that automate data movement and governance tasks at large scale.

- **Cyber recovery orchestration:** The "Mediaflux Point-in-Time" capability functions as a versioned file system that can revert an entire namespace to a known good state, providing rapid restoration in the event of an attack.
- **AI data provenance and lineage:** Arcitecta offers an advanced lineage model designed to trace data processing histories and dependencies, enabling reproducibility and integrity verification in research and AI environments.

Arcitecta was classified as an Outperformer due to its accelerated innovation pace and ongoing expansion of the Mediaflux platform into a full-featured metadata-driven data fabric. Over the past year, the company has delivered significant enhancements in automation, cyber recovery workflow, and AI-assisted metadata management, supported by frequent software releases and a strong engineering execution culture. Its roadmap emphasizes advanced metadata analytics, integrated AI pipelines, and expanded connectivity to structured data sources, positioning Arcitecta ahead of many competitors in technical depth and release velocity.

OPPORTUNITIES

Arcitecta has room for improvement in a few decision criteria, including:

- **Data security posture management (DSPM):** While its security model is strong on access control and auditing, it falls short of a full DSPM solution, lacking native capabilities for discovering data exposure risks or modeling an attack's "blast radius."
- **Generative AI data curation:** The solution provides a solid foundation for this feature with its native vector database. To advance to an exceptional level, Arcitecta could develop a closed-loop feedback mechanism from LLMs to autonomously refine and rank the curated datasets for maximum relevance.
- **Unified data fabric extension:** The platform's XODB provides a capable foundation for managing diverse data types. To improve, Arcitecta could develop robust, out-of-the-box connectors and query federation tools for common structured enterprise databases, moving beyond its strong HPC/research focus.

PURCHASE CONSIDERATIONS

Arcitecta's licensing is a key differentiator, priced by user count with no capacity-based fees, offering highly predictable and transparent pricing for large-scale environments. It is licensed as a complete, integrated software fabric. This approach is designed for large enterprises, particularly in data-intensive verticals, and is not a point solution. Deployment is customer-managed, and the platform's power and complexity mean professional services are highly recommended

for architectural design, integration, and workflow development. The complexity is a trade-off for its exceptional scalability and customizability.

USE CASES

Mediaflux supports a wide range of use cases, but it is specifically optimized for data-intensive industry verticals. It is a strong fit for high-performance computing (HPC), life sciences (genomics, research), defense and intelligence, and media and entertainment sectors, for which managing exabyte-scale data, complex workflows, and data orchestration are primary drivers.

BigID: BigID Next

SOLUTION OVERVIEW

BigID is an intelligence-led software vendor providing an AI-powered platform for data discovery, security, and governance. Its flagship solution, BigID Next, operates as a unified control plane with a modular "app framework," allowing customers to flexibly license capabilities for DSPM, privacy, and AI compliance. The platform acts as a storage-agnostic intelligence layer spanning the entire data estate, including structured databases, unstructured files, and multicloud environments. By decoupling intelligence from the underlying infrastructure, BigID provides a comprehensive view of data risk without requiring data movement or rehydration.

BigID is positioned as a Leader and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

BigID scored well on a number of decision criteria, including:

- **Data security posture management (DSPM):** The platform offers a comprehensive DSPM approach that discovers and classifies sensitive data, maps permissions, and identifies exposure risks such as over-permissioned repositories.
- **AI-driven data intelligence:** The solution applies AI and ML to discover and contextualize enterprise data across large, heterogeneous environments, improving visibility and decision-making.
- **Customizable AI/ML classification:** The offering includes a flexible, extensible classification engine that supports advanced ML, natural language processing (NLP), and computer-vision models, allowing teams to design and refine custom classifiers for specialized data types.

OPPORTUNITIES

BigID has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** The current solution focuses on assessing data exposure and attack impact but lacks integrated recovery workflows such as immutable snapshot management and automated restoration.
- **Automated policy enforcement:** BigID uses its extensive metadata insights to drive robust policy orchestration. For lifecycle management, the platform automates governance workflows by triggering tiering, archival, and retention actions across connected storage systems, effectively bridging the gap between discovery and remediation.
- **AI data provenance and lineage:** Lineage tracing could be enhanced with finer-grained, automated visualization of data movement through AI/ML pipelines to improve auditability and explainability.

PURCHASE CONSIDERATIONS

BigID is licensed via a subscription model based on the modules (apps) activated and the amount of data managed. This modularity can create complexity in determining the right SKU bundle. End users buy it for its deep intelligence capabilities to be deployed in a best-of-breed ecosystem alongside existing storage and backup solutions. The platform is highly flexible, offering SaaS, customer-managed (on-prem), customer-hosted cloud deployment, and hybrid deployment options (pairing SaaS management with local pods). Given its power and the scale of data it connects to, professional services are common for initial deployment and tuning.

USE CASES

BigID is adopted to support specific, high-value, data-centric use cases. Key purchase drivers include DSPM, sensitive data discovery for privacy and compliance (GDPR, CCPA), data minimization (ROT reduction), and enabling AI governance by providing a unified, classified inventory of enterprise data for AI/ML teams.

Cohesity: Cohesity Data Cloud

SOLUTION OVERVIEW

Cohesity is a data security and management vendor focused on consolidating secondary data, data protection, and data security onto a single, unified architecture. The Cohesity Data Cloud, managed via the Helios SaaS plane, is a single platform-centric solution combining data protection (DataProtect), file/object services (SmartFiles), data security/DSPM, and GenAI (Cohesity Gaia). Cohesity's approach provides a broad, generalized

solution for holistically managing and securing data. The recent merger with Veritas signals a significant expansion of its market footprint. Cohesity delivers an aggressive roadmap, evident from its rapid expansion into AI and security, and values rapid advancement and frequent releases and updates.

Cohesity is positioned as a Leader and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar report.

STRENGTHS

Cohesity scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The platform delivers a comprehensive recovery workflow centered on Cohesity FortKnox for immutable, air-gapped data vaulting and RecoveryAgent for automated clean room recovery. This integrated approach combines anomaly detection, secure vaulting, and orchestrated restoration blueprints within a single management plane, streamlining response to ransomware incidents.
- **Automated policy enforcement:** An SLA-based policy engine is exceptional, automating not just data protection and retention but also immutability, cyber vaulting, and data lifecycle management from a single, unified framework.
- **Data security posture management (DSPM):** Cohesity utilizes a native AI/ML-based classification engine to discover sensitive data and correlate it with recovery points. To extend this visibility, the Security Center Marketplace integrates with dedicated DSPM platforms (such as Cyera) to ingest deep risk insights, enabling organizations to align protection policies, detect compliance gaps, and remediate ROT data from a unified dashboard.

OPPORTUNITIES

Cohesity has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** Cohesity has significantly expanded its intelligence capabilities with Gaia (RAG) and Copilot, which allow users to query backup content for business context. However, because these insights are derived from ingested snapshots, they reflect the state of the data at the last protection interval. Expanding this analysis to live, in-place data activity (prior to backup ingestion) remains an opportunity to provide real-time operational awareness comparable to primary storage analytics.
- **Customizable AI/ML classification:** The system relies mainly on predefined classification libraries. Adding user-friendly tools for training and validating custom ML models would increase flexibility.
- **Unified data fabric extension:** While Cohesity effectively unifies the storage and cataloging of diverse workloads (files, objects, backups) on a single platform, performing deep cross-domain

analysis (for example, correlating unstructured content with structured database records) typically relies on external integrations (such as Snowflake). Developing broader native analytical capabilities to perform these correlations in place (without requiring data movement to third-party warehouses) would further enhance the platform's value as a unified data fabric.

PURCHASE CONSIDERATIONS

The Cohesity Data Cloud operates on a term-based subscription model, typically priced by capacity (per backend terabyte). Licensing is streamlined through tiered bundles, such as Enterprise and Premium editions, which now include advanced capabilities like data classification and threat protection that were previously sold separately. Specialized services like Cyber Vaulting (FortKnox) can be added to these bundles. A key strength is the platform's deployment flexibility, allowing customers to choose between physical appliances, virtual software, or a fully SaaS-hosted model, all managed via the unified Helios control plane. This "software-defined" approach allows organizations to modernize legacy backup infrastructure without being locked into a rigid hardware form factor.

USE CASES

Cohesity supports a broad, horizontal set of use cases across most industries, especially finance, healthcare, and public sector. The primary purchase drivers are modernizing backup and data protection, consolidating NAS infrastructure (via SmartFiles), and, most critically, adopting a unified platform for cyber resilience, including ransomware detection and automated recovery.

CTERA: CTERA Enterprise File Services Platform

SOLUTION OVERVIEW

CTERA delivers a unified file and data-services platform that extends across edge, core, and cloud environments through the CTERA Portal management plane and Edge Filers. This architecture creates a single global namespace that centralizes collaboration, lifecycle policies, and security while maintaining high-performance access for remote users. Evolving beyond infrastructure modernization, the platform now functions as a hybrid cloud orchestration layer integrating file services, ransomware resilience, AI-driven content intelligence, and governance under one control model.

CTERA is positioned as a Challenger and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

CTERA scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The platform employs CTERA Ransom Protect, which combines immutable snapshotting to object storage with AI-based anomaly detection. This enables rapid, orchestrated rollback of the file system to a clean, pre-attack state.
- **Automated policy enforcement:** A centralized policy framework automates snapshot scheduling, retention, and caching behavior across distributed filers, ensuring consistent governance without manual intervention at the edge.
- **Data security posture management (DSPM):** The Data Discovery module provides baseline functionality for identifying sensitive data and detecting unusual access behavior across edge and cloud environments.

OPPORTUNITIES

CTERA has room for improvement in a few decision criteria, including:

- **Enterprise workflow integration:** CTERA has strong integrations with key security and backup vendors (e.g., Varonis, Cohesity). However, expanding its library of pre-built API connectors for broader IT service management (ITSM) and downstream analytics tools would further enhance its versatility as a data hub.
- **Customizable AI/ML classification:** The platform currently relies on user-defined JSON schemas to guide AI extraction, rather than offering trainable machine-learning models (e.g., "teach-by-example"). While this schema-based approach allows flexible custom tagging, it lacks the simplicity and computational efficiency of dedicated, user-trainable classifiers for identifying proprietary data patterns at scale.
- **Generative AI data curation:** While the platform successfully delivers native RAG pipelines, expanding the ecosystem of supported vector databases and downstream LLM integrations beyond the current defaults would allow customers greater flexibility in building bespoke AI applications.

PURCHASE CONSIDERATIONS

CTERA is purpose-built as a cyber-resilient global file system that unifies edge, core, and cloud. Licensing is subscription-based (capacity or per-user), with the option to add the Data Intelligence module for advanced observability and AI readiness. While it supports air-gapped private deployments ("Private Cloud IT") for sovereign use cases, its modern value proposition increasingly leverages the CTERA Insight SaaS layer to provide native data curation and RAG pipeline integration. This makes it a dual-use platform: a robust infrastructure layer for distributed file storage and a governed knowledge source for enterprise AI models.

USE CASES

The platform is adopted as a unified hybrid cloud file and data services layer that centralizes management across edge, core, and cloud environments. While the primary driver remains modernizing distributed NAS infrastructure to deliver a single, global namespace for ROBO collaboration, the use case has expanded to include data observability and AI readiness. Organizations utilize the platform's integrated analytics for forensic usage tracking and capacity planning while leveraging its data intelligence capabilities to curate and feed unstructured content securely into generative AI (RAG) pipelines.

Data Dynamics: Zubin

SOLUTION OVERVIEW

Data Dynamics is an intelligence-led software vendor that recently consolidated its portfolio into Zubin, a single, unified UDM platform. Zubin is a storage-agnostic, low-code solution that provides in-place discovery, classification, security, and governance across hybrid and multicloud environments. Its "pragmatic AI" approach leverages metadata inferencing, sampling, and anomaly detection to manage data at petabyte scale. This provides a focus acting as a unified intelligence and governance layer on top of an organization's existing, heterogeneous storage infrastructure. As a new platform launched in 2024, Zubin has an aggressive quarterly release cadence and a roadmap focused on rapid advancement in AI and security.

Data Dynamics is positioned as a Challenger and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Data Dynamics scored well on a number of decision criteria, including:

- **AI-driven data intelligence:** The platform uses pragmatic AI methods such as metadata inferencing and natural language processing to contextualize unstructured data and connect it with relevant business entities.
- **Data security posture management (DSPM):** A zero trust security framework integrates sensitive-data discovery and permissions analysis, providing continuous insight into exposure risks and facilitating policy-based remediation.
- **Automated policy enforcement:** The rules-based governance engine supports automated lifecycle, compliance, and cost-optimization actions, allowing administrators to operationalize policies efficiently.

OPPORTUNITIES

Data Dynamics has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** The current design emphasizes anomaly detection and identification of threats but lacks native recovery orchestration, such as immutable snapshot handling or automated restoration workflows.
- **Customizable AI/ML classification:** Classification models can be extended, but low-code tools for training and testing custom ML models would simplify adoption.
- **Unified data fabric extension:** The solution's correlation between unstructured and structured data remains limited. Developing native connectors for database integration would support full data-fabric functionality.

PURCHASE CONSIDERATIONS

Licensing for the Zubin platform is subscription-based, with the solution offered as a single, integrated product rather than a complex set of modules. The end users license Zubin as a storage-agnostic intelligence layer to be used in combination with their existing multi-vendor storage solutions. Its "low-code, self-service" design aims to simplify deployment and management compared to more complex platforms. Data Dynamics' heritage in data mobility also makes it a strong consideration for organizations with complex migration needs, as this capability is integral to the platform.

USE CASES

The platform is adopted as a unified intelligence and governance layer for managing unstructured data across hybrid and multicloud environments. Organizations use the solution to centralize discovery, classification, movement, and policy enforcement under a single control plane that supports compliance, optimization, and AI readiness. Typical deployments include enterprise-wide risk mitigation through integrated DSPM and sensitive-data discovery, automated cost control via ROT data management, and data mobility workflows that migrate and tier content seamlessly between platforms. The platform also plays a key role in preparing curated, trusted, and business-relevant datasets forGenAI applications, integrating governance, lifecycle management, and intelligence into a cohesive framework.

Dell Technologies: PowerScale, ObjectScale

SOLUTION OVERVIEW

Dell Technologies is an infrastructure-led vendor whose UDM solution is a comprehensive portfolio anchored by its PowerScale scale-out network attached storage (NAS), ObjectScale enterprise object storage, and PowerScale for Public Cloud. This storage foundation is augmented by an integrated software-defined suite, including the PowerScale Cybersecurity Suite (featuring Ransomware Defender), MetadataIQ, and the new Dell AI Data Platform (which integrates a data search engine and a data processing engine powered by Apache Spark). This portfolio constitutes a broad, generalized platform play approach, leveraging its market-leading infrastructure base as a foundation for advanced cyber recovery and AI data services.

Dell Technologies is positioned as a Challenger and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Dell Technologies scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The PowerScale Cybersecurity Suite, combined with an operational air gap vault, provides an end-to-end, automated workflow for identifying threats, protecting data, and executing a rapid, clean recovery.
- **Generative AI data curation:** The AI Data Platform introduces a dedicated RAG Connector that analyzes metadata to accelerate data preparation and ingestion for GenAI models, enhancing integration between storage intelligence and AI pipelines.
- **Unified data fabric extension:** The Dell AI Data Platform distinguishes itself with a dual-engine architecture: a Data Analytics Engine based on Trino (powered by Starburst) for federated querying, and a Data Processing Engine powered by Apache Spark for heavy-duty data engineering. Notably, Dell has optimized these engines with the NVIDIA Warp framework, delivering unique hardware-accelerated performance for data simulation and processing tasks, an integration currently exclusive to the Dell ecosystem.

OPPORTUNITIES

Dell Technologies has room for improvement in a few decision criteria, including:

- **Customizable AI/ML classification:** Current search tools provide basic, built-in classification, but the solution lacks a user-facing toolkit that customers can use to easily train new, custom ML models based on their own document examples.

- **AI-driven data intelligence:** The Elastic-based data search engine offers strong metadata search but would benefit from enhanced contextual analytics to infer relationships and semantic meaning across datasets automatically.
- **AI data provenance and lineage:** Native tools for depicting lineage and tracking the provenance of data used in AI training are still emerging. Enhancements here would improve transparency and auditability.

PURCHASE CONSIDERATIONS

Licensing for Dell's UDM solution is portfolio-based, offering flexibility to consume PowerScale and ObjectScale infrastructure via traditional CapEx or as a subscription through Dell APEX. However, this modularity can introduce complexity, as advanced capabilities, such as the Cybersecurity Suite and the AI Data Platform, are licensed as add-ons to the core storage. While PowerScale OneFS remains a mature and relatively simple platform to manage at scale, the comprehensive UDM solution (integrating Elastic, Starburst, and Apache Spark) adds architectural weight, typically necessitating professional services for optimal design and integration.

USE CASES

Dell supports a broad, horizontal set of use cases across most industries. Key purchase drivers include high-performance file storage for demanding workloads (AI/GenAI, media and entertainment, life sciences), consolidating legacy NAS infrastructure, and adopting a comprehensive, infrastructure-integrated cyber recovery solution.

Druva: Druva Data Security Cloud

SOLUTION OVERVIEW

Druva is a cloud-native vendor that delivers its UDM capabilities through the Druva Data Security Cloud, a 100% SaaS platform. This unified platform consolidates data protection, cyber recovery, and data governance by analyzing data it ingests from endpoints, data centers, and SaaS applications (such as M365) for backup. This provides a broad, holistic platform for data resilience rather than acting as primary storage. As a SaaS-native platform, Druva delivers an aggressive roadmap, evidenced by its comprehensive DruAI suite. Built on the Metagraph metadata foundation, DruAI includes specialized agents for ransomware recovery and environment optimization, as well as DruAssist for support, embedding GenAI deeply into the operational workflow.

Druva is positioned as a Leader and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Druva scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The architecture combines anomaly detection with automated, policy-driven recovery workflows, enabling rapid data restoration via immutable snapshots and verified, clean recovery points.
- **Automated policy enforcement:** The SLA-driven policy engine automates backup, retention, replication, and archival across hybrid and cloud environments. Integrated with Druva's global metadata layer, it reduces administrative overhead through centralized, profile-based management in the SaaS console.
- **AI-driven data intelligence:** Druva differentiates itself with DruAI, an agentic AI engine built on the Dru MetaGraph. This architecture moves the platform beyond static reporting by actively analyzing distributed metadata to predict risks, troubleshoot failures, and automate complex management tasks via natural language, effectively turning backup data into an operational knowledge base.

OPPORTUNITIES

Druva has room for improvement in a few decision criteria, including:

- **Data security posture management (DSPM):** While Dru MetaGraph provides exceptional visibility into data risks across the backup estate, the platform operates primarily on ingested snapshots. Expanding into full DSPM would involve tighter integration with primary storage systems to enable real-time changes to access controls and remediation at the source (before the next backup cycle), thereby closing the latency gap, compared to in-place, event-driven security tools.
- **Customizable AI/ML classification:** While Druva offers verifiable "Curated Snapshots" for security and ransomware recovery (removing malware to create clean restore points), the platform does not yet provide workflows to cleanse, vectorize, or expose these backup datasets for external GenAI training. The current curation capabilities focus on operational hygiene rather than preparing a semantic knowledge corpus for customer-built RAG applications.
- **Unified data fabric extension:** While Druva effectively centralizes diverse workloads (including endpoints, SaaS apps, and cloud-native data) for resilience, the platform focuses on backup datasets. It does not provide a native data fabric to correlate unstructured file content with structured production data records (such as joining file metadata with active database rows) for unified business intelligence.

PURCHASE CONSIDERATIONS

Licensing for the Druva Data Security Cloud follows a 100% SaaS model, with transparent, consumption-based pricing aligned to workload type (typically per user for Microsoft 365 and endpoints, or per terabyte of source data for servers and NAS). Notably, this model includes storage, egress, and support, minimizing the hidden cloud costs often associated with hyperscale deployments. Druva is a best-of-breed solution designed to replace legacy backup infrastructure, leveraging backup data to deliver advanced cyber resilience and governance. Its greatest strength lies in operational simplicity: as a fully managed cloud platform, it requires no customer infrastructure, patching, or upgrades, enabling rapid deployment and policy automation across hybrid environments.

USE CASES

Druva is adopted to support specific, high-value use cases. The primary purchase drivers are modernizing data protection, adopting a cloud-first strategy, and gaining a comprehensive, automated ransomware detection and recovery solution. It is also strong in e-discovery and compliance for protected data (including M365) and securing data on endpoints.

Hammerspace: Hammerspace Data Platform

SOLUTION OVERVIEW

Hammerspace is an intelligence-led software vendor that provides the Hammerspace Data Platform, a parallel global file system. This software-defined platform is storage-agnostic, unifying disparate on-prem (NFS, S3-enabled) and cloud storage silos into a single, high-performance, metadata-driven namespace. This architecture provides a focus designed to provide data orchestration and nondisruptive data mobility for specific, data-intensive workloads. Hammerspace delivers an aggressive roadmap for its unique architecture and is highly flexible and responsive to the evolving demands of HPC and AI, valuing rapid advancement.

Hammerspace is positioned as a Leader and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Hammerspace scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The platform includes a metadata-driven orchestration engine to automatically enforce complex policies for data placement, tiering, and mobility across the entire global fabric.

- **Cyber recovery orchestration:** The architecture leverages global file-system snapshots to enable rapid, orchestrated recovery, allowing administrators to nondisruptively rewind the namespace to a pre-attack state.
- **AI data provenance and lineage:** A metadata-centric design supports lineage tracking, providing the foundation for auditability and reproducibility in AI/ML workflows.

OPPORTUNITIES

Hammerspace has room for improvement in a few decision criteria, including:

- **Data security posture management (DSPM):** The solution currently lacks integrated sensitive data discovery, exposure mapping, and other content-aware security features needed for comprehensive data risk visibility.
- **AI-driven data intelligence:** Existing analytics focus on infrastructure orchestration and data movement, with no native capabilities to examine content or infer business relationships from file data.
- **Customizable AI/ML classification:** Configuration options are limited to metadata extensions, and there is no user-facing toolkit for training or deploying ML models to classify data based on content.

PURCHASE CONSIDERATIONS

Licensing for Hammerspace is subscription-based, typically priced by the volume of data being orchestrated. While fundamentally a software-defined global file system, procurement is flexible. Organizations can deploy it as software on existing infrastructure, purchase it as an integrated hardware appliance (via partners like Supermicro and Hitachi Vantara), or run it as a fully cloud-native instance. This versatility supports diverse infrastructure strategies, from pure on-premises to hybrid and multicloud. However, given the sophistication of establishing a global namespace across these heterogeneous environments, professional services are often recommended for architectural design and policy optimization.

USE CASES

The platform is widely adopted as a global data-orchestration layer for hybrid and multicloud environments, serving industries such as high-performance computing (HPC), AI/ML, media and entertainment, and life sciences. Organizations use the solution to unify previously siloed storage systems into a single, metadata-driven control plane that automates data placement, mobility, and lifecycle management across edge, core, and cloud resources. Key use cases include enabling automated data pipelines feeding AI and ML model training,

optimizing high-performance workloads through intelligent data locality controls, and enforcing consistent security and compliance policies globally. The platform also supports modern data fabric and hybrid cloud initiatives for which automated orchestration and real-time data access are critical to achieving scalability and business agility.

Hitachi Vantara: Virtual Storage Platform One (VSP One) and Hitachi Content Platform (HCP)

SOLUTION OVERVIEW

Hitachi Vantara is an infrastructure-led vendor whose UDM solution is a comprehensive, multiproduct portfolio managed under the VSP 360 control plane. The foundation consists of its storage platforms, including VSP One (File/Object) and Hitachi Content Platform (HCP). UDM intelligence, classification, and workflow capabilities are delivered by Hitachi Content Intelligence (HCI) and the Pentaho data platform. This integrated portfolio leverages enterprise-grade infrastructure as a foundation for advanced governance and analytics, specifically preparing data for AI workloads on Hitachi iQ.

Hitachi Vantara is positioned as a Leader and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Hitachi Vantara scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The platform combines the policy automation and classification capabilities of HCI and Pentaho with immutable object lock, retention, and tiering policies across VSP One file and object services. This delivers comprehensive lifecycle control across hybrid environments, with VSP 360 enforcing and orchestrating policy execution and providing auditable visibility into policy actions across the platform.
- **Enterprise workflow integration:** The Pentaho data integration and orchestration environment, together with HCI, connects to a broad set of enterprise systems. This powerful combination enables automated workflows that move, index, and transform data across diverse sources with minimal manual intervention, bridging the gap between structured operations and unstructured content governance.
- **Unified data fabric extension:** Pentaho and HCI connect to and operate on structured and unstructured data across internal and external sources. Underlying this, the VSP One platform provides the governed file and object services where unstructured data is persistently stored, protected, and policy-enforced, enabling a unified, auditable view across distributed environments.

OPPORTUNITIES

Hitachi Vantara has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** Existing capabilities such as HCP immutability and integration with CyberSense provide a strong foundation; however, unifying these into a single, fully automated recovery workflow would simplify cyber resilience operations.
- **Data security posture management (DSPM):** Discovery and classification features in HCI could be strengthened by incorporating access control insight and threat detection analytics within a centralized DSPM dashboard for guided remediation.
- **AI-driven data intelligence:** Current AI and analytics functionality in HCI and Pentaho operate as separate modules. Greater convergence of these components would enable more automated inference of business context and relationships.

PURCHASE CONSIDERATIONS

Hitachi Vantara is streamlining its licensing through the VSP One strategy, which consolidates previously disparate components (block, file, object) into a unified hybrid cloud data platform. While the solution offers immense breadth, integrating advanced storage (HCP), compute (HCI), and analytics (Pentaho), the new model emphasizes consumption-based subscriptions (EverFlex) and SaaS-managed operations to simplify procurement. This evolution addresses the historical complexity of deploying individually licensed products, offering large enterprises a scalable, flexible foundation that is easier to consume while retaining the power of a bespoke infrastructure solution.

USE CASES

Hitachi Vantara supports horizontal use cases across most industries, with particular strength in highly regulated sectors such as finance, healthcare, and government. Key purchase drivers include enterprise-scale governance and compliance (including Cohasset-validated immutability and retention), modernizing data analytics pipelines (via Pentaho and Hitachi iQ), and building curated data lakes for AI and GenAI workloads. These solutions leverage the integrated VSP One platform, with centralized operational visibility and control provided through VSP 360.

HPE: HPE GreenLake

SOLUTION OVERVIEW

Hewlett Packard Enterprise is an infrastructure-led vendor whose UDM solution is a comprehensive portfolio delivered via its HPE GreenLake hybrid cloud platform. This is not a single product but an integrated set of services, including HPE GreenLake for File Storage, HPE Object Storage, the HPE Ezmeral Data Fabric (for AI/analytics), and HPE GreenLake for Cyber Recovery. This portfolio constitutes a broad approach, leveraging its high-performance infrastructure as a foundation for advanced AI/ML and cyber resilience workloads. HPE delivers an aggressive roadmap focused on integrating its portfolio under the GreenLake platform and values rapid advancement, especially in AI and security.

HPE is positioned as a Challenger and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

HPE scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The HPE GreenLake for Cyber Recovery solution is a purpose-built, automated, air-gapped vaulting and recovery orchestration platform, providing an exceptional end-to-end workflow.
- **Unified data fabric extension:** The Ezmeral Data Fabric offers a mature framework designed to build a unified namespace and catalog that integrates both unstructured (file/object) and structured (database) sources to support AI pipelines.
- **Enterprise workflow integration:** The Ezmeral platform integrates deeply with data-science tools, job schedulers, and analytics systems, enabling seamless orchestration of MLOps, data pipelines, and cross-platform automation.

OPPORTUNITIES

HPE has room for improvement in a few decision criteria, including:

- **Data security posture management (DSPM):** The portfolio currently lacks content-aware capabilities for sensitive-data discovery, exposure mapping, and risk remediation, leaving customers reliant on third-party integrations for these functions.
- **AI-driven data intelligence:** Current focus is on infrastructure (AIOps) and pipelines (MLOps), but the solution lacks the native UDM features to analyze file content and autonomously infer business context or semantic relationships for governance.
- **Customizable AI/ML classification:** While Ezmeral is a platform for building ML models, it does not offer a simple, built-in, user-facing trainable classifier for tagging files, which is a key capability for modern UDM.

PURCHASE CONSIDERATIONS

Licensing for HPE's UDM solution is portfolio-based and integrated into the HPE GreenLake consumption model, offering pay-as-you-go flexibility that can simplify procurement, though portfolio complexity remains. This is an integrated solution for large enterprises. The solution's power, particularly the HPE Ezmeral component, makes it inherently complex and requires significant data engineering and IT expertise to deploy, integrate, and manage. It is not a simple "general IT admin" solution, and professional services are typically required.

USE CASES

HPE's solution is horizontal but particularly strong in specific, high-value use cases. Key purchase drivers include building advanced AI/ML and GenAI data pipelines (via Ezmeral), adopting a comprehensive, automated cyber recovery solution (via GreenLake for Cyber Recovery), and infrastructure modernization for mission-critical file and object data.

IBM*

SOLUTION OVERVIEW

IBM is an infrastructure-led vendor evolving rapidly through innovation within a well-established technology foundation. Its unstructured data management strategy combines established infrastructure with new, integrated data- and AI-driven advancements across the IBM Storage Scale (high-performance file) and IBM Storage Defender (cyber recovery), IBM Security Guardium (DSPM), and the watsonx AI and data platform. Together, these components form a cohesive architecture that unifies protection, intelligence, and analytics for hybrid cloud deployments.

IBM is positioned as a Leader and Fast Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

IBM scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** Storage Defender provides an integrated recovery framework that combines anomaly detection, cyber vaulting, and automated clean-recovery processes, enabling rapid and validated restoration following a security incident.
- **Generative AI data curation:** The watsonx.data platform automates data ingest, tagging, and quality assessment while maintaining lineage integrity, allowing organizations to govern and prepare data efficiently for AI and RAG pipelines.

- **Unified data fabric extension:** The watsonx.data lakehouse architecture enables native joins between structured and unstructured datasets, creating a single analytic and governance layer across hybrid environments for unified insight.

OPPORTUNITIES

IBM has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** Existing functionality could be enhanced through automated contextual inference and relationship mapping across unstructured sources, reducing the need for manual metadata preparation.
- **Customizable AI/ML classification:** While watsonx.ai offers sophisticated ML training tools, IBM could strengthen this feature by embedding easier low-code capabilities within the UDM workflow to let users train and deploy organization-specific classifiers directly.
- **Conversational data intelligence:** Conversational access through watsonx Assistant is available separately. Closer native integration within the UDM environment would allow users to query and manage data more intuitively.

PURCHASE CONSIDERATIONS

Licensing for IBM's UDM capabilities is generally portfolio-based, with individual products (Storage Scale, Guardium, and watsonx) licensed separately under flexible enterprise or consumption agreements. This model provides maximum scalability and customization but adds pricing and SKU complexity. The solution is designed for large enterprises and heavily regulated industries. Deployment typically requires professional services to architect and integrate the various components. IBM's long-term support model and hybrid design provide strong ROI for large-scale environments that demand enterprise-grade resilience, compliance, and AI integration.

USE CASES

IBM's platform supports broad, horizontal use across industries including finance, healthcare, government, and research. Key purchase drivers include enterprise-scale cyber resilience, compliance enforcement and auditability, large-scale data governance, and AI-driven data curation using watsonx for analytics and model training workloads.

Komprise

Komprise is an intelligence-led software vendor that delivers a UDM platform built on its Transparent Move Technology (TMT) architecture. The solution provides in-place analytics, automated data movement, governance, and deep integrations with major storage and cloud providers. Komprise analyzes file and object data across heterogeneous environments without disrupting user access, offering granular policy automation and analytics via its "Deep Analytics" and "Smart Data Workflows" features. The platform focuses on hybrid and multicloud data management for governance, cost optimization, and analytics enablement.

Komprise is positioned as a Leader and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Komprise scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The distributed architecture and Transparent Move Technology enable exceptional, massively parallel policy automation for tiering, archiving, and ROT data deletion across heterogeneous storage environments.
- **AI-driven data intelligence:** The Deep Analytics engine provides superior intelligence, inferring context such as project, owner, and usage patterns to drive data movement and cost optimization decisions automatically.
- **Generative AI data curation:** Smart Data Workflows streamline dataset preparation for AI and RAG pipelines by automating tagging, filtering, and movement of relevant files to cloud or AI environments for model training and inference.

OPPORTUNITIES

Komprise has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** Komprise effectively shrinks the attack surface by offloading cold data to immutable, air-gapped storage (Object Lock). The platform currently focuses on prevention and isolation rather than active restoration. Expanding integration with dedicated backup and cyber recovery platforms (to coordinate the rapid restore of hot data alongside the relinking of cold data) would allow Komprise to participate more fully in the end-to-end recovery lifecycle rather than just the isolation phase.
- **Data security posture management (DSPM):** Komprise has significantly improved its security capabilities with Smart Data Workflows, which now enable automated remediation (tiering, confinement, or exclusion) of sensitive data found during scans. To evolve into a comprehensive DSPM platform, the opportunity remains to expand into identity and entitlement analysis, mapping

user access permissions (ACLs) against data sensitivity to identify overprivileged users and open shares, rather than focusing solely on the data's location.

- **Customizable AI/ML classification:** While the platform supports query-based tagging, adding a user-facing toolkit for training and validating ML classifiers based on sample documents would improve adaptability for specialized domains.

PURCHASE CONSIDERATIONS

Komprise is licensed via a subscription model typically based on managed capacity or number of file operations. It is designed as an overlay intelligence and management service deployed alongside existing storage systems and cloud environments. The SaaS-based management console simplifies deployment, requiring minimal infrastructure, though professional services are offered for policy design and analytics tuning in large-scale deployments.

USE CASES

Komprise is deployed across industries seeking hybrid visibility and automation, with particular strength in healthcare and highly regulated sectors that require governed AI adoption. Key purchase drivers include global metadatabase cataloging (via the Global File Index), petabyte-scale cost optimization, and Intelligent AI Ingestion, which filters noise and eliminates ROT to prepare high-quality datasets. A critical differentiator is the platform's storage-agnostic, lock-in-free architecture, which ensures mobilized data remains natively accessible across all silos without requiring proprietary agents or stubs.

Nasuni: Nasuni File Data Platform

Nasuni is an infrastructure-led vendor providing a global file system solution designed for hybrid and multicloud data management. The Nasuni File Data Platform, powered by the UniFS global file system, consolidates file storage, backup, and collaboration under a single namespace with centralized management. Data is stored as objects in major public clouds (AWS, Azure, GCP) or private cloud object stores for example, Cloudian and IBM COS), while frequently accessed files are cached locally at edge sites for optimal performance. The platform integrates natively with Nasuni Edge Appliances, Nasuni Orchestration Center, File IQ Premium, and Nasuni Advanced Web Access to deliver secure, simplified file access and visibility at scale.

Nasuni is positioned as a Challenger and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Nasuni scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The platform incorporates native, immutable snapshots and rapid, global recovery workflows capable of restoring entire file systems to a verified pre-attack state, providing strong ransomware resilience.
- **Automated policy enforcement:** The cloud-backed architecture enables policy-driven snapshot frequency, retention, caching, and tiering rules. These automated policies simplify data lifecycle management and reduce administrative overhead.
- **Enterprise workflow integration:** Integration with key security, monitoring, and collaboration ecosystems, such as SIEM/SOAR tools and Microsoft 365, enables coordinated visibility and workflow automation for distributed teams.

OPPORTUNITIES

Nasuni has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** The platform currently lacks native AI-based discovery and context inference. Expanding autonomous metadata analysis would extend functionality beyond infrastructure management toward true data intelligence.
- **Generative AI data curation:** With the release of the Nasuni Data Service (NDS), the platform now natively exposes file data to GenAI and RAG pipelines via direct object access. The next step in this evolution would be to develop deeper native curation capabilities (such as automated vectorization or semantic chunking) to reduce reliance on external ecosystem partners for dataset preparation.
- **Customizable AI/ML classification:** Although existing integrations provide analytics support, an embedded, trainable ML classification toolkit would enhance the ability to identify organization-specific data patterns across the global file system.

PURCHASE CONSIDERATIONS

Nasuni uses a subscription-based licensing model typically priced by capacity and the number of managed edge sites. The platform is delivered as a cloud-native service supported by lightweight edge appliances that connect to centralized object storage, minimizing infrastructure overhead and simplifying global administration. Nasuni provides an integrated global file and data services layer that unifies file storage, data protection, and governance across distributed environments. Professional services support large-scale deployments, focusing on global name space design, migration planning, policy automation, and integration with enterprise security and analytics platforms.

USE CASES

Nasuni supports hybrid and multisite enterprises in sectors such as architecture, engineering, manufacturing, media, and finance. Typical use cases include global NAS consolidation, ransomware resilience with rapid recovery, remote collaboration enablement through a unified file system, and simplified hybrid cloud data management. With the addition of File IQ Premium, the platform is also increasingly adopted for intelligent data insights, enabling organizations to analyze file metadata and activity trends for governance, capacity planning, and forensic auditing.

NetApp: BlueXP Data Sense*

NetApp is a platform-centric vendor delivering an integrated UDM solution under its BlueXP management plane. BlueXP unifies visibility, policy management, and orchestration across ONTAP (file/block), StorageGRID (object), and public cloud storage in AWS, Azure, and GCP. The platform combines data mobility, protection, security, and analytics within a single control plane. NetApp's Cloud Insights, Data Sense, and Cloud Sync services extend the ecosystem, enabling classification, cost optimization, and governance of hybrid and multicloud environments.

NetApp is positioned as a Leader and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

NetApp scored well on a number of decision criteria, including:

- **Automated policy enforcement:** BlueXP and ONTAP provide a unified policy framework that automates data-tiering, protection, and compliance operations consistently across on-prem and cloud environments.
- **Data security posture management (DSPM):** BlueXP Data Sense delivers superior DSPM capabilities, discovering sensitive data, mapping risk exposure, and enforcing corrective or protective policies through integration with ONTAP security controls.
- **Cyber recovery orchestration:** The integration of SnapLock, SnapVault, and SnapMirror technologies provides a robust cyber recovery workflow with immutable snapshots, isolated vaulting, and orchestrated restoration—critical for ransomware resilience.

OPPORTUNITIES

NetApp has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** BlueXP Data Sense provides strong metadata analytics, but the platform could advance further by introducing semantic inference and contextual relationship mapping across unstructured content to enhance autonomous insight generation.

- **Generative AI data curation:** Existing data pipeline capabilities could be extended to include a dedicated, automated RAG workflow for preparing, versioning, and exposing curated datasets directly to AI platforms and enterprise LLMs.
- **Customizable AI/ML classification:** Classification capabilities rely primarily on predefined patterns. Adding low-code tools for training custom ML classifiers would help customers target proprietary or domain-specific data types more effectively.

PURCHASE CONSIDERATIONS

NetApp offers flexible subscription and capacity-based licensing across the BlueXP platform, with usage metering unified through its digital marketplace. Deployment options span customer-managed, SaaS, and hybrid cloud configurations. While ideal for enterprises seeking strategic hybrid cloud consistency, integration of advanced services (Data Sense, Cloud Insights) may introduce additional SKUs and configuration complexity. Professional services or partners can assist with governance onboarding and large-scale automation policy design.

USE CASES

NetApp serves a broad range of industries, including financial services, manufacturing, and the public sector. Key use cases include hybrid data protection with immutable snapshots and automated ransomware recovery, multicloud governance and compliance automation, AI/analytics-ready data lake optimization, and unified policy control across hybrid environments.

Panzura: CloudFS*

Panzura is an infrastructure-led vendor providing the Panzura CloudFS, a global file system designed to consolidate distributed file storage, facilitate collaboration, and deliver cyber-resilient data management. Managed through Panzura Data Services (PDS), the platform unifies file access, search, analytics, and recovery across on-prem and public cloud object storage. It focuses on security, ransomware recovery, and simplified operations for distributed enterprises. The architecture combines cloud-backed immutability, centralized intelligence, and edge performance.

Panzura is positioned as a Challenger and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Panzura scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The platform includes immutable snapshots and orchestrated recovery capabilities through Panzura Data Services, enabling quick restoration of file environments following a ransomware or data-corruption event.
- **Automated policy enforcement:** The platform employs centralized policy automation for file lifecycle management, replication, and access control, ensuring consistent data governance across distributed global file systems.
- **Enterprise workflow integration:** It integrates with enterprise monitoring and governance ecosystems, including SIEM and backup tools, to extend data protection intelligence across connected infrastructure.

OPPORTUNITIES

Panzura has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** AI capabilities remain focused on operational telemetry. Expansion into content analysis and business context inference would strengthen its analytics value.
- **Generative AI data curation:** While native dataset preparation for AI and RAG pipelines is not yet automated, the platform's metadata-driven global file system provides the foundation for identifying, tagging, and retrieving relevant datasets for AI applications. Enhancing these base capabilities with end-to-end orchestration tools would make the solution more fully aligned with enterprise AI workflows.
- **Customizable AI/ML classification:** Classification is largely metadata-based. Introducing trainable, ML-driven classification models would enhance precision for organization-specific data types.

PURCHASE CONSIDERATIONS

Panzura licenses its solutions via subscription, typically based on storage capacity or active edge nodes. The architecture is customer-managed, deployed as a global file system and SaaS-delivered management interface. Within a hybrid environment, it often complements existing storage systems and backup solutions. Deployment is straightforward for distributed environments, though larger rollouts may benefit from professional services for namespace design and policy alignment.

USE CASES

The platform is broadly adopted as a unified file and data management foundation that supports distributed, hybrid cloud environments. Deployments often replace legacy NAS and distributed file system (DFS) infrastructures to deliver a single, cloud-orchestrated namespace and centralized data services spanning performance, security, and governance. Organizations use the solution to consolidate file workloads, simplify multisite collaboration, and ensure cyber resilience through immutable snapshots and orchestrated recovery. It also underpins hybrid cloud initiatives, enabling consistent data access, synchronization, and lifecycle automation across on-prem and cloud resources. Adoption is common in globally distributed industries such as engineering, media and entertainment, and architecture, for which shared high-performance data workflows and integrated management capabilities are essential.

Pure Storage

Pure Storage is an infrastructure vendor delivering a unified data platform built around FlashBlade, FlashArray, and the Pure Fusion management plane. Together, these components provide a high-performance, scale-out architecture for unstructured data workloads, AI pipelines, and cyber-resilient operations. The Pure1 cloud-based management engine unifies these systems through predictive analytics, policy automation, and cost optimization intelligence. Recent advancements focus on deeper AI integration, autonomous performance tuning, and hybrid cloud connectivity, extending the platform beyond traditional storage modernization into intelligent data services that simplify management and accelerate data-driven innovation.

Pure Storage is positioned as a Challenger and Forward Mover in the Innovation/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Pure Storage scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** SafeMode Snapshots provide immutable data protection with automated recovery workflows that enable fast, reliable restoration across on-prem arrays and connected cloud targets.
- **Automated policy enforcement:** The Pure Fusion control plane (integrated into Pure1) utilizes presets to automate policy enforcement across the fleet. By defining storage classes and compliance rules once, administrators can ensure consistent protection, retention, snapshot scheduling, and service-level compliance at scale without manual intervention on individual arrays.

- **Enterprise workflow integration:** Pure1 integrates with a broad ecosystem of enterprise and security tools such as ServiceNow, Splunk, and vSphere, allowing streamlined alerting, orchestration, and automation across IT environments.

OPPORTUNITIES

Pure Storage has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** Current capabilities focus on telemetry and system performance. Adding content-aware analytics that infer business context from unstructured file data would extend AI insight beyond infrastructure operations.
- **Generative AI data curation:** While Flash supports high-performance access for AI workloads, integrated pipelines for curation, versioning, and management of GenAI datasets would enhance the solution's utility in AI data engineering.
- **Customizable AI/ML classification:** The platform offers robust telemetry analytics but lacks a user-facing ML training and classification interface for tagging and identifying organization-specific data types consistently across the environment.

Pure Storage is classified as a Forward Mover, reflecting reliable execution in data resilience and ease of use but a cautious approach to advancing AI-driven data services. While the company continues to excel in operational simplicity, its recent innovations—such as Pure Fusion (utilizing presets for fleet-wide policy automation) and Pure1 AI Copilot—remain primarily focused on extending infrastructure telemetry and management rather than introducing groundbreaking content intelligence. This measured pace reinforces its leadership in storage automation but suggests a roadmap prioritized around platform consistency, positioning Pure Storage as a deliberate mover within an increasingly AI-accelerated market.

PURCHASE CONSIDERATIONS

Pure Storage licenses its FlashBlade and FlashArray platforms through capacity-based or Evergreen subscription models, providing predictable lifetime ownership. The vendor has significantly simplified fleet operations by embedding the Pure Fusion control plane directly into the Purity OS, allowing customers to deploy an Enterprise Data Cloud architecture without additional licensing. Pure Fusion's preset functionality is a key differentiator, enabling automated, policy-driven storage compliance and governance at scale for example, standardizing protection and QoS rules). While Pure Fusion handles infrastructure governance, organizations may still need third-party tools for granular content governance (like PII classification) within files themselves.

USE CASES

Pure Storage's platform supports a wide range of enterprise and data-intensive workloads across sectors such as enterprise IT, media and entertainment, healthcare, financial services, and AI research. Typical deployments extend beyond standalone storage modernization to deliver unified data resilience, analytics, and AI enablement within a single management framework. Common use cases include consolidating unstructured data under a centralized control plane, building ransomware-resilient recovery architectures, powering hybrid analytics and GenAI pipelines, and modernizing on-prem and cloud infrastructure for consistent performance and low-latency access.

Quantum

Quantum is an infrastructure-focused vendor specializing in unstructured data storage and lifecycle management for performance-intensive, content-rich environments. Its UDM portfolio includes Quantum Myriad (next-generation scale-out file and object storage), ActiveScale (an object storage platform with built-in immutability and a cold storage tier), and CatDV (an AI-enabled media asset management system). Together, they deliver a hybrid architecture focused on data resilience, media-aware workflows, and long-term archival storage. Quantum's solutions are widely used across media, entertainment, research, and government sectors where data scale, preservation, and security are critical.

Quantum is positioned as a Challenger and Fast Mover in the Maturity/Platform Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Quantum scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** ActiveScale ObjectLock, combined with cyber.vaulting integration, delivers immutable data protection and a structured workflow that enables rapid ransomware recovery without data loss.
- **Automated policy enforcement:** Rule-based lifecycle management automatically handles archiving, tiering, and replication across Myriad, ActiveScale, and tape libraries, ensuring policy consistency throughout the data estate
- **Unified data fabric extension:** Through CatDV's workflow orchestration, Quantum effectively bridges unstructured file systems with workflows such as content indexing, version management, and AI analysis, empowering broader data insight across ecosystems.

OPPORTUNITIES

Quantum has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** While CatDV provides workflow connectors to multiple AI services (for example, Rekognition, TwelveLabs) and supports local AI transcription, Quantum could further enhance this by integrating deeper contextual inference and semantic analysis directly across Myriad and ActiveScale to unify metadata intelligence enterprise-wide.
- **Generative AI data curation:** Native pipelines for curating and exposing content for RAG or other GenAI workloads are not yet available. Building this bridge among storage, metadata, and AI platforms would enhance strategic relevance.
- **Customizable AI/ML classification:** Existing classification depends on pretrained AI models. Introducing an intuitive interface for training organization-specific models would increase flexibility and domain adaptability.

PURCHASE CONSIDERATIONS

Quantum offers perpetual and subscription-based licensing across its portfolio, with models typically tied to capacity or module selection (like for Myriad, ActiveScale, CatDV). The solutions are customer-managed, reflecting Quantum's infrastructure focus. While deployment can be complex for multisite or hybrid environments, extensive professional services and integration support are available. Quantum's strengths lie in its durability, cost efficiency, and media workflow optimization rather than in holistic data intelligence.

USE CASES

Quantum's platform serves as a comprehensive foundation for hybrid data management across unstructured, structured, and media-rich environments. Typical deployments span industries such as media and entertainment, life sciences, aerospace, and government, for which scalable, metadata-aware infrastructure is essential for content management and analytics. Key use cases include active content production pipelines, hybrid data fabric deployments supporting AI and ML workloads, unified protection and recovery across storage tiers, and long-term digital preservation requiring immutable, high-performance storage integrated under a single orchestration layer.

Qumulo: Qumulo Core File Data Platform

Qumulo is an infrastructure-led vendor whose core offering, the Qumulo Core File Data Platform, provides a modern, scale-out file and object storage architecture for unstructured data at enterprise scale. The platform is software-defined and runs identically across on-prem environments and public clouds (AWS, Azure, GCP, and OCI), enabling seamless hybrid operations. Qumulo's strength lies in analytics-driven visibility, data mobility, and high performance for file-based workloads such as AI/ML, energy, media production, financial services, healthcare imaging, life sciences, public sector, higher education, and research. The solution combines real-time observability, policy automation, and integration with modern data pipelines.

Qumulo is positioned as a Challenger and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Qumulo scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The software-defined file system provides policy-driven automation for data placement and retention, allowing administrators to manage tiering and lifecycle transitions between active, archive, and cloud tiers with minimal configuration.
- **Cyber recovery orchestration:** The architecture supports rapid detection, protection, and recovery from ransomware through immutable snapshots, anomaly detection, and clean recovery options across hybrid deployments.
- **Enterprise workflow integration:** The platform integrates natively with major enterprise ecosystems, including Splunk, NVIDIA DGX, and cloud-native data services. These integrations enable hybrid workflows for AI, data analytics, and security events.

OPPORTUNITIES

Qumulo has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** While Qumulo excels at high-performance data delivery for AI model training (feeding GPUs), the platform relied primarily on external downstream services for content analysis during the review period. Deeper, intrinsic AI capabilities to derive business context or semantic relationships directly within the file system were not yet fully integrated compared to competitors with native metadata intelligence engines.
- **Generative AI data curation:** The architecture is widely used within AI pipelines, yet it lacks native automation for dataset curation, tagging, and preparation, which are increasingly essential for GenAI and RAG applications.

- **Customizable AI/ML classification:** While Qumulo Helios utilizes ML for infrastructure telemetry (predicting capacity, performance issues, and anomalies), the platform lacks user-facing, trainable classifiers to analyze file content. Developing tools to identify proprietary document types or business-specific data patterns (beyond standard metadata) would expand the solution's value for governance and compliance workloads.

PURCHASE CONSIDERATIONS

Qumulo is licensed via subscription and available as software-defined storage on customer-managed infrastructure, as well as SaaS-managed deployments in AWS, Azure, GCP, or OCI. It provides straightforward consumption pricing scaled by usable capacity. Designed as a complete file data platform, it typically replaces legacy NAS or parallel file systems. Qumulo's modern API-first design minimizes operational overhead, though large-scale hybrid environments may require professional services for capacity planning and performance tuning.

USE CASES

Qumulo serves high-performance, data-intensive industries such as life sciences, media and entertainment, healthcare, financial services, public sector, and research. Key purchase drivers include scalable file data consolidation, hybrid AI/ML and analytics enablement, ransomware-resilient recovery, and simplified hybrid management of petabyte-scale storage environments.

Rubrik: Rubrik Security Cloud (RSC)

Rubrik is an intelligence-led data security and resilience platform provider. Its flagship solution, the Rubrik Security Cloud (RSC), unifies backup, recovery, DSPM, and data threat analytics into a single control plane. Built on a zero trust architecture, the platform integrates deeply with workloads across on-prem, SaaS, and public cloud environments. Rubrik's strategy focuses on using AI and analytics to deliver autonomous risk reduction, continuous compliance, and seamless cyber recovery while preparing data for broader AI/ML consumption.

Rubrik is positioned as a Leader and Outperformer in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Rubrik scored well on a number of decision criteria, including:

- **Cyber recovery orchestration:** The platform delivers an end-to-end recovery workflow that combines immutable snapshots, precise blast radius identification, and automated clean room restoration, supported by integrated vaulting and reporting that ensure a verifiable recovery SLA.

- **Data security posture management (DSPM):** A unified DSPM framework identifies sensitive data, evaluates access exposure, and analyzes data flows across the environment. This module integrates directly with Rubrik Radar to correlate sensitivity with threat activity, highlighting potential risks and automating remediation actions to reduce the blast radius of attacks.
- **Automated policy enforcement:** The intelligent policy engine applies SLA-based automation for protection, archival, immutability, and compliance controls, maintaining consistent governance across hybrid and multicloud environments.

Rubrik is classified as an Outperformer because of its rapid innovation pace, strong release cadence, expanding AI integrations, and clear strategic direction toward autonomous data security and GenAI readiness. The company's accelerated development in AI-driven posture management and cyber recovery distinctly positions it ahead of competitors in the short to mid-term.

OPPORTUNITIES

Rubrik has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** Current analytics center on security telemetry. Expanding functionality to include autonomous context mapping and unstructured data insight generation would broaden the platform's analytical depth.
- **Generative AI data curation:** While Rubrik Security Cloud secures and indexes vast volumes of data, it currently lacks a native RAG pipeline that leverages this curated dataset directly for AI/ML workloads.
- **Customizable AI/ML classification:** The classification engine is strong for sensitive data discovery but lacks user-facing tools to train new ML models for domain-specific document types or unstructured data categories.

PURCHASE CONSIDERATIONS

Rubrik Security Cloud is licensed by capacity and feature tier via SaaS and subscription models. Its unified management console simplifies deployment and ongoing operations with low administrative overhead. As a comprehensive platform, Rubrik often replaces multiple point products for backup, archival actions, cyber recovery, and DSPM. The company's strong enterprise integrations (ServiceNow, Splunk, Microsoft Sentinel) enhance value for customers seeking automation across security and IT ecosystems.

USE CASES

Rubrik serves midsize to large enterprises across regulated sectors such as financial services, healthcare, and government, as well as unregulated industries seeking advanced data protection. Typical use cases include ransomware resilience and incident recovery, data security posture management, zero trust data protection, compliance automation, and preparing trusted, protected datasets for emerging AI and analytics initiatives.

Scality: RING, ARTESCA*

Scality is an infrastructure-led vendor specializing in hybrid and cloud-scale unstructured data storage. Its flagship products, Scality RING and ARTESCA, form a comprehensive, software-defined storage platform supporting file and object workloads at petabyte scale. RING provides large-scale, enterprise-grade durability and immutability for core and archival data, while ARTESCA extends lightweight, cloud-native object services for edge and AI/ML environments. The portfolio integrates policy-based automation, data lifecycle management, and multitenant orchestration for hybrid cloud deployments.

Scality is positioned as a Challenger and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Scality scored well on a number of decision criteria, including:

- **Automated policy enforcement:** The integrated rules engine automates tiering, replication, and immutability enforcement across on-prem and object storage tiers, allowing administrators to execute policies at scale with minimal manual management.
- **Cyber recovery orchestration:** Object-lock immutability, WORM capabilities, and geo-distributed erasure coding together provide strong ransomware resilience and data integrity assurance, enabling faster recovery after cyber events.
- **Enterprise workflow integration:** RING and ARTESCA expose comprehensive RESTful APIs and support major enterprise platforms (like ServiceNow and Splunk) through native connectors, enabling bidirectional event-driven orchestration and simplified hybrid integration.

OPPORTUNITIES

Scality has room for improvement in a few decision criteria, including:

- **AI-driven data intelligence:** Intelligence features remain focused on infrastructure performance and metadata reporting. Adding semantic analysis and contextual inference within ARTESCA would broaden enterprise data visibility.

- **Generative AI data curation:** While the object storage platforms provide durability and scale, native pipelines for curating and exposing datasets for RAG and GenAI workloads are not yet available. Building direct integration with AI frameworks would increase strategic value.
- **Customizable AI/ML classification:** Metadata tagging is limited to predefined models. Embedding user-trainable ML classification would enhance content recognition and compliance-driven insight generation.

PURCHASE CONSIDERATIONS

Scality licenses RING and ARTESCA as software-defined solutions on a subscription or capacity-based model. RING is typically deployed on-prem or within private clouds, and ARTESCA offers a lightweight, cloud-native edge deployment model. The architecture suits customers seeking high-durability, hybrid data persistence over general-purpose intelligence. Professional services are often used for architectural design, scaling, and integration into enterprise automation frameworks.

USE CASES

Scality serves large enterprises, service providers, and government agencies managing petabyte-scale unstructured data. Core use cases include immutable backup repositories, AI/ML data lake storage, media content archives, and long-term digital preservation requiring enterprise-grade durability and hybrid cloud flexibility.

StrongLink*

StrongLink is an intelligence-led data management platform vendor offering a unified, metadata-driven solution designed to automate the discovery, movement, and preservation of unstructured data across hybrid storage environments. Its core platform, StrongLink, functions as a self-optimizing data orchestration engine built on a powerful metadata catalog that aggregates, enriches, and correlates information from all connected file, object, and tape systems. StrongLink provides automated lifecycle management, intelligent policy enforcement, and deep integrations for analytics, compliance, and cyber resilience.

StrongLink is positioned as a Leader and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

StrongLink scored well on a number of decision criteria, including:

- **AI-driven data intelligence:** The metadata-driven AI engine automatically discovers relationships between files, infers context, and builds an adaptive "data knowledge graph," improving data visibility and operational insight across environments.

- **Automated policy enforcement:** The platform provides a superior rules engine that autonomously executes complex multistage lifecycle policies spanning movement, archival action, and replication based on metadata triggers and real-time system analytics.
- **Enterprise workflow integration:** The open, API-first architecture integrates natively with diverse enterprise platforms and automation frameworks, enabling bidirectional orchestration with SIEM, SOAR, and ITSM systems.

OPPORTUNITIES

StrongLink has room for improvement in a few decision criteria, including:

- **Generative AI data curation:** Metadata intelligence creates a solid foundation for RAG workflows, but native automation for curating, versioning, and securely exposing datasets to GenAI tools remains undeveloped. Proactive integration would enhance relevance for emerging AI applications.
- **Customizable AI/ML classification:** While StrongLink's AI tags data relationships autonomously, it does not yet offer a simple, end user-facing interface for training custom ML classifiers or importing externally developed models for domain-specific categorization.
- **AI data provenance and lineage:** The platform tracks metadata lineage but could expand toward full dataset lineage visualization and automated auditing, ensuring transparent traceability of data transformations across hybrid and multicloud environments.

PURCHASE CONSIDERATIONS

StrongLink's enterprise licensing model is typically subscription-based and scaled by capacity or number of managed endpoints. The solution is deployed on-prem or in hybrid configurations, requiring minimal storage vendor lock-in thanks to its hardware-agnostic design. Professional services are usually recommended for metadata model customization, integration with enterprise automation systems, and advanced policy configuration across heterogeneous environments.

USE CASES

StrongLink is adopted by large enterprises and government organizations seeking unified, intelligent data control across multivendor storage ecosystems. Key use cases include metadata-driven lifecycle management, automated file-to-object tiering and archiving, hybrid cloud orchestration, AI-ready data preparation, and compliance-driven data preservation.

Varonis: Varonis Data Security Platform (DSP)

Varonis is an AI-powered data security platform providing comprehensive visibility, classification, remediation at scale, and risk management for enterprise data. The Varonis Data Security Platform (DSP) integrates file analysis, activity monitoring, and automated policy enforcement across on-prem file systems, NAS, and cloud collaboration platforms, including Microsoft 365, Salesforce, ServiceNow, Snowflake, Databricks, and Slack. Its unified architecture detects and mitigates data risk by combining metadata analytics, UEBA, and DSPM capabilities to identify, prioritize, and remediate exposure while ensuring compliance. Additionally, Varonis provides dedicated security controls for AI adoption, securing data usage in Microsoft 365 Copilot, ChatGPT, and Salesforce Agentforce.

Varonis is positioned as a Challenger and Fast Mover in the Innovation/Feature Play quadrant of the unstructured data management Radar chart.

STRENGTHS

Varonis scored well on a number of decision criteria, including:

- **Data security posture management (DSPM):** The platform provides an advanced DSPM framework that continuously maps sensitive data exposure, permissions, and data flows while automating guided remediation across enterprise file and SaaS environments.
- **AI-driven data intelligence:** ML models infer data context automatically, identifying sensitive information, linking it to user behavior, and correlating activities across systems to enable precise risk analytics and insight.
- **Automated policy enforcement:** The policy engine enforces access and retention controls and autonomously revokes risky permissions or resolves noncompliance issues across hybrid environments.

OPPORTUNITIES

Varonis has room for improvement in a few decision criteria, including:

- **Cyber recovery orchestration:** While Varonis can identify compromises and help isolate affected data, it lacks integrated recovery orchestration workflows such as vaulting, clean room restore, and automatic reintegration with backup systems.
- **Generative AI data curation:** Current focus areas center on security and risk analytics. Developing capabilities to curate and safely expose anonymized datasets for AI and RAG pipelines would broaden its applicability for enterprise AI initiatives.
- **Customizable AI/ML classification:** Prebuilt AI classifiers cover common data types, but introducing low-code for training domain-specific models would enhance customization and extend classification precision for specialized use cases.

PURCHASE CONSIDERATIONS

Varonis is offered via subscription licensing, primarily priced by user count, which provides cost predictability regardless of data growth. The notable exception is for IaaS coverage (for example, AWS S3, Azure Blob), which is typically metered by data volume. The platform is deployed as a SaaS solution with optional self-hosted collectors for hybrid integrations. It is often adopted to complement existing backup, storage, and governance systems while serving as the enterprise's primary data security and risk management layer. The solution's rich analytics and automation depth make it enterprise-ready but best suited for organizations with dedicated security or compliance teams.

USE CASES

Varonis is chosen across regulated and security-sensitive industries, including finance, healthcare, energy, and government. Key use cases include DSPM and least-privilege enforcement, insider threat detection, compliance reporting (GDPR, HIPAA, CCPA), ransomware detection and containment, and risk analytics for sensitive unstructured data in on-prem and cloud collaboration environments.

The industry's pivot from traditional infrastructure management to contextual data intelligence reflects an overarching truth: unstructured data has become the substrate upon which AI, security, and business transformation converge.

06 Analyst's Outlook

THE UNSTRUCTURED DATA MANAGEMENT (UDM) market has entered a period of accelerated evolution, defined not only by scale but by intelligence. Once centered on storage efficiency and visibility, the category now sits at the intersection of AI readiness, cyber resilience, and compliance automation. Modern enterprises are awash in unindexed information: emails, design files, digital media, and machine data that represent both operational risk and untapped strategic value. The industry's pivot from traditional infrastructure management to contextual data intelligence reflects an overarching truth: unstructured data has become the substrate upon which AI, security, and business transformation converge.

From a purchaser's perspective, this is not a homogeneous market. Solutions can be broadly grouped into infrastructure-led, platform-centric, and intelligence-led categories. Infrastructure vendors such as Dell Technologies, Hitachi Vantara, HPE, and Pure Storage focus on data durability, cyber recovery, and performance at scale. Platform providers, including NetApp and IBM, have formed integrated suites that combine hybrid data control with AI and compliance frameworks. Intelligence-led players, such as BigID, Komprise, and Varonis, distinguish themselves by transforming unstructured data into governed, contextualized, and actionable insights. Understanding these layers provides the foundation for a meaningful evaluation. Buyers must determine whether their primary pain point is operational efficiency, security, or business enablement and then align solutions accordingly.

Across the market, three themes dominate purchase decisions. First, data security posture management (DSPM) and cyber resilience are now Key Features, triggered by a relentless surge in ransomware and privacy mandates. Platforms integrating immutable architectures, automated remediation, and AI-driven threat analysis are setting the functional baseline for enterprise-grade adoption. Second, AI enablement, particularly GenAI data curation, has emerged as a key differentiator. Vendors able to bridge governance and AI pipelines are being fast-tracked in enterprise transformation programs. Finally, the concept of a unified data fabric is reshaping strategic procurement. The value no longer lies in mass centralization but in metadata-driven unification, allowing organizations to operate as if data were centralized without physically moving it.

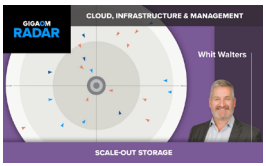
For IT decision-makers, the next best action is to begin with a comprehensive audit of their existing unstructured data estate. Map not only the places where data lives but who accesses it and how policies are applied. This inventory serves as the basis for prioritizing control objectives, including risk reduction, compliance, and optimization. Enterprises should pilot one or two use cases first, such as automated data tiering or DSPM deployment, before scaling into AI-driven classification or unified governance. Cross-functional collaboration among storage, security, and data science teams is critical; silos will impede both visibility

and AI readiness. Buyers should also vet vendor claims regarding AI capabilities, specifically, whether models are embedded and can infer context autonomously or whether they are simply rebranded metadata search tools.

Looking forward, the UDM market is converging with adjacent domains, including data security, AI governance, and observability. Over the next 12 to 24 months, expect the most innovation to involve three categories: self-optimizing data fabrics that blend structured and unstructured sources, continuous compliance models that operate autonomously, and generative data curation engines that refine enterprise training sets in real time. The winners will be the vendors that integrate seamlessly across ecosystems rather than maintain bounded silos. For practitioners, the objective is clear: treat UDM as a foundational capability for the AI era, not a housekeeping function. Organizations that modernize now by combining governance rigor with AI-powered intelligence will position themselves not only to survive data growth but also to turn it into a strategic advantage.

To learn about related topics in this space, check out the following GigaOm Radar reports:

GigaOm Radar for **Enterprise Object Storage v5.0**



GigaOm Radar for **Scale-Out Storage v6.0**

GigaOm Radar for **Primary Storage v5.0**

Methodology

Hitachi Vantara

*Vendors marked with an asterisk did not participate in our research process for the Radar report, and their capsules and scoring were compiled via desk research.

For more information about our research process for Radar reports, please visit our Methodology.

About James Brown

James is a seasoned tech leader in data analytics, storage, AI/ML, and strategic partnerships, excelling in scalability, deployment efficiency, customer satisfaction, and contributing to industry knowledge through blogs and whitepapers.

GIGAOM

About GigaOm

GigaOm provides technical, operational, and business advice for IT's strategic digital enterprise and business initiatives. Enterprise business leaders, CIOs, and technology organizations partner with GigaOm for practical, actionable, strategic, and visionary advice for modernizing and transforming their business. GigaOm's advice empowers enterprises to successfully compete in an increasingly complicated business atmosphere that requires a solid understanding of constantly changing customer demands.

GigaOm works directly with enterprises both inside and outside of the IT organization to apply proven research and methodologies designed to avoid pitfalls and roadblocks while balancing risk and innovation. Research methodologies include but are not limited to adoption and benchmarking surveys, use cases, interviews, ROI/TCO, market landscapes, strategic trends, and technical benchmarks. Our analysts possess 20+ years of experience advising a spectrum of clients from early adopters to mainstream enterprises.

GigaOm's perspective is that of the unbiased enterprise practitioner. Through this perspective, GigaOm connects with engaged and loyal subscribers on a deep and meaningful level.

GIGAOM

GigaOm is built on the perspective of professionals who have operated, deployed, and managed the technologies we assess. Our practitioner lens brings nuance, realism, and foresight, ensuring our research is trusted, actionable, and aligned to the real pressures facing modern enterprises.