BEDROCK SECURITY

The Power of Metadata Lakes:
Transforming Data Security and Governance

Introduction

Cloud. SaaS. AI. Every corner of the enterprise is generating, moving, and modifying data faster than security teams can keep up. The result? A sprawling, high-velocity environment where data is everywhere, and context is nowhere.

According to **Bedrock's 2025 Enterprise Data Security Confidence Index**, **82%** of security leaders say they still can't reliably find or classify data across production, customer, and employee environments. Over half admit they lack continuous visibility, needing days, or weeks, to locate sensitive assets when it matters most.

That's not just a tooling problem. It's a metadata problem.

In fragmented, multi-cloud environments, traditional governance models break. Security teams default to generic controls, reactive policies, and superficial checks. This leaves critical gaps: misclassified PII, broken access controls, stale policies, all waiting to be exploited.

A metadata lake changes the game. It consolidates metadata, lineage, access, classification, and more, across cloud, SaaS, and infrastructure. It creates a shared, real-time knowledge layer that connects engineering, security, and governance teams around a single source of truth.

And the impact is clear:

- **84%** of security leaders said a metadata lake gives them accurate data inventory across systems
- **78%** cited stronger signal across security tools
- **75%** saw sharper detection and prevention through context-aware policy enforcement

This paper explores how metadata lakes work, why they've become essential in modern cloud and AI security architecture, and how they're helping enterprises shift from reactive controls to proactive, intelligent data governance.

BEDROCK SECURITY

Understanding Technical & Business Metadata

Metadata, often described as “data about data,” provides insights into the structure, classification, and usage of information within your organization. It enables you to understand where your data resides, how it is used, and who has access to it. This information is crucial for managing security policies, ensuring compliance with regulatory requirements, and optimizing data governance. Metadata is categorized into two primary types: technical metadata and business metadata.

Technical Metadata

Technical metadata includes information such as data schemas, structures, transfer protocols, and storage locations. It describes how data is organized, processed, and shared across enterprise systems.

Business Metadata

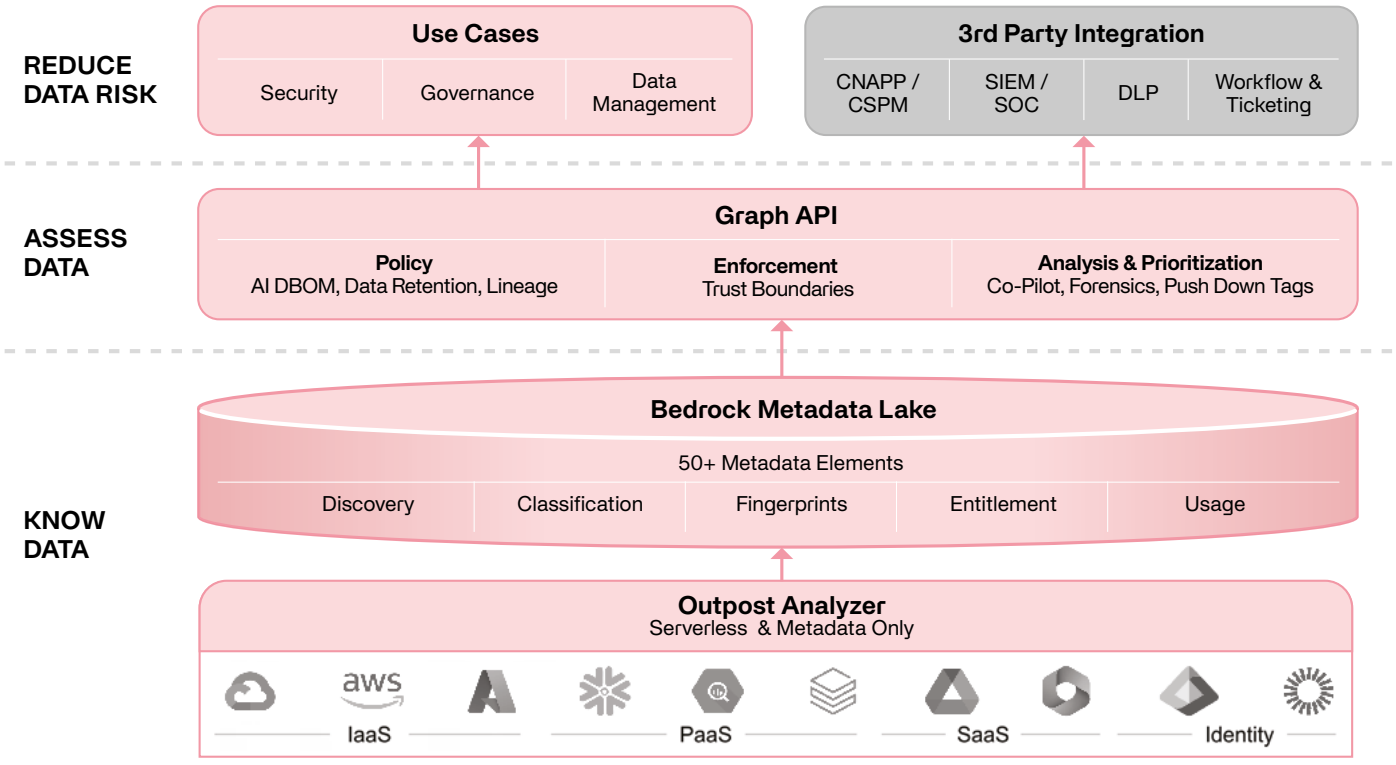
Business metadata encompasses details about data ownership, governance policies, regulatory requirements, and classification labels. When used effectively, metadata enables enterprises to gain deeper visibility into their data assets, improve classification and governance, and optimize data usage for security and compliance purposes.

By leveraging both types of metadata strategically, your security team can learn what data is most important to the business, enforce access control policies on that data, detect anomalies, and streamline security operations while maintaining transparency in data management.

The Concept of a Metadata Lake

A metadata lake is a centralized repository that aggregates metadata from multiple sources, allowing you to manage security, governance, and compliance in a scalable and efficient manner. Unlike traditional data lakes, which focus on storing raw data, a metadata lake is designed to provide rich context and insights into data assets. This ensures faster and more accurate data management, security, and GRC without reading or storing your actual data, which causes additional data security challenges.

The Bedrock Security Advantage

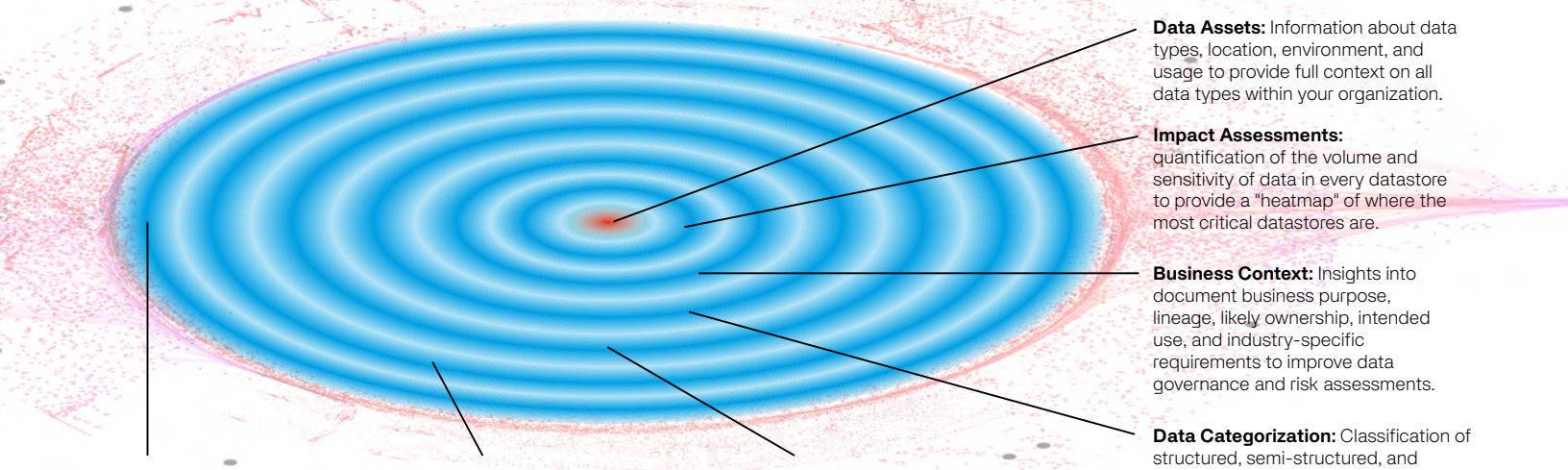


Bedrock Security has revolutionized data security by operationalizing metadata lakes, moving beyond traditional DSPM limitations. Unlike standard DSPM solutions that rely on static, manual classification and exported reports requiring extensive manual synthesis, Bedrock's metadata lake provides dynamic, AI-powered classification and continuous sensitivity tagging. Its intuitive natural-language query capabilities enable security teams to gain immediate, actionable insights instead of manually analyzing reports. By offering real-time visibility and continuously updated context, Bedrock ensures your security posture adapts swiftly to changing threats, achieving unmatched responsiveness and accuracy across your entire data landscape.

Key Components of a Metadata Lake

Metadata Information (What is Stored)

Metadata lakes store a broad range of information about your data assets, allowing your security, data management, and GRC teams to gain deep visibility into how data is structured, categorized, and used. Some of the key metadata elements include information on:



Data Assets: Information about data types, location, environment, and usage to provide full context on all data types within your organization.

Impact Assessments: quantification of the volume and sensitivity of data in every datastore to provide a "heatmap" of where the most critical datastores are.

Business Context: Insights into document business purpose, lineage, likely ownership, intended use, and industry-specific requirements to improve data governance and risk assessments.

Data Categorization: Classification of structured, semi-structured, and unstructured data for better visibility and improved security controls, along with identification of data subject to regulation (e.g. PII, PHI), corporate intellectual property, and other sensitive data types that require protection under compliance frameworks.

Data Jurisdiction: Identifying where the subjects of the data reside to ensure compliance with regional data residency requirements such as GDPR, CCPA, and industry-specific mandates like HIPAA and PCI-DSS.

Data Access & Usage: Monitoring who has access to data, how frequently it is used, and whether access patterns align with your security policies.

Service Vulnerabilities: Identifying security risks associated with your deployed cloud services and their access to data, helping to mitigate threats before they become breaches.

Metadata Collection (How Data is Gathered & Used)

To be effective, a metadata lake must continuously collect, analyze, and share metadata across security tools and data governance platforms. Methods for metadata collection include:

- **Enterprise-wide Data Discovery:** Identifying and cataloging all data sources across your SaaS, IaaS, and PaaS environments to build a comprehensive metadata inventory.
- **Data Boundary Mapping:** Establishing security policies based on metadata analysis, allowing you to define protection rules dynamically without requiring extensive manual effort.
- **Full Entitlement Chain Analysis:** Monitoring access permissions across multiple platforms and identity providers to ensure least-privilege access and prevent data overexposure.
- **API-Enabled Security Integrations:** Enhancing security tools like data loss prevention (DLP), cloud-native application protection platforms (CNAPP), and cloud security posture management (CSPM) with metadata-driven insights.
- **Scalable, Serverless Architecture:** Ensuring cost-effective, low-latency metadata processing that can scale with data growth and evolving security requirements.

Use Cases for a Metadata Lake

A metadata lake serves as a centralized repository that aggregates and manages metadata across your organization's data landscape, enabling several critical use cases, in addition to traditional DSPM capabilities such as:

- **Sensitive Data Discovery** Identify and classify sensitive data across cloud and on-prem environments.
- **Data Risk Assessment** Evaluate data exposure risks based on sensitivity, access permissions, and usage patterns.
- **Access and Entitlement Review** Analyze who has access to what data and enforce least privilege policies.
- **Data Movement Tracking** Monitor data flows to detect unauthorized transfers or shadow data copies.
- **Policy Enforcement and Alerting** Automate enforcement of data security policies and trigger alerts for violations.
- **Compliance Reporting** Generate reports to demonstrate adherence to regulations like GDPR, HIPAA, and CCPA.
- **Data Retention and Minimization** Ensure unnecessary or outdated sensitive data is identified and safely removed.
- **Shadow Data Detection** Surface unmanaged or forgotten data assets that pose security or compliance risks.

Additional Bedrock metadata lake use cases include:

Natural-Language Queries & Continuous Responses

Unlike traditional DSPM solutions that typically export static reports requiring manual analysis, Bedrock Security's metadata lake empowers you to answer complex, plain-language queries instantly through a copilot function. Non-technical users can easily obtain insights into data access patterns, permissions, and usage without needing SQL expertise, enabling swift, informed decision-making and significantly reducing response times.

Dynamic, No-Rules AI-Driven Classification & Tagging

Bedrock Security's metadata lake continuously updates sensitivity tags using advanced AI that dynamically adapts to changes in your data environment. Traditional DSPM tools often rely on static rules or manual processes, leading to outdated or inaccurate classifications. With a metadata lake, classifications remain current, reflecting continuous changes in data context, lineage, and business usage, ensuring sensitive information is always accurately identified and protected.

The Power of Metadata Lakes: Transforming Data Security and Governance

BEDROCK SECURITY

Forensics & Vulnerability Breach Impact Analysis

When a breach or vulnerability is discovered, Bedrock Security's metadata lake enables rapid impact analysis by automatically mapping all downstream data dependencies. This accelerates incident response by revealing the full blast radius, identifying compromised systems, and reducing investigation time, without requiring manual data tracing or reliance on incomplete lineage records.

Lineage Analysis for Governance & Compliance

Bedrock Security's metadata lake provides complete, automated visibility into the lifecycle of data across your environment. By tracing data from origin to consumption, it ensures integrity, detects unauthorized use, and simplifies compliance reporting. This lineage clarity supports better governance, risk management, and audit-readiness—without the complexity of fragmented or siloed tools.

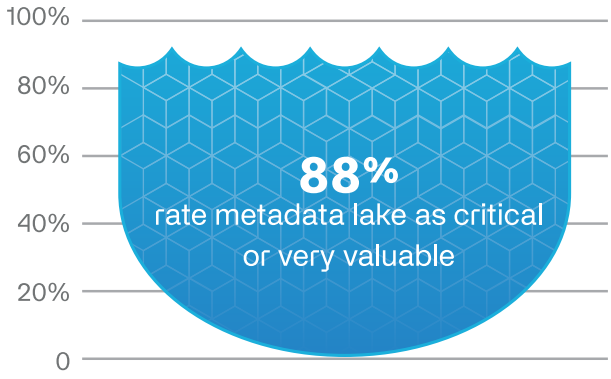
Cross-Cloud Metadata Integration

Traditional DSPM solutions typically manage metadata in isolated silos. In contrast, Bedrock Security's metadata lake integrates metadata across security, governance, compliance, and data management domains. This comprehensive integration eliminates visibility gaps, provides holistic insights, and ensures that policies and controls are consistently applied across your entire data infrastructure.

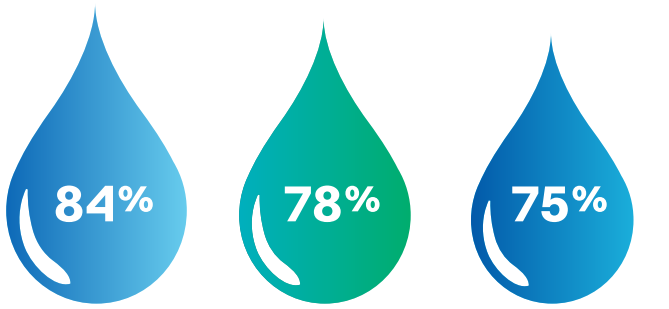
Flexible Entitlement Management & Verification

DSPM tools typically provide rigid, role-based access controls that require frequent manual updates. Bedrock Security's metadata lake continuously verifies and dynamically adjusts entitlements based on real-time metadata insights and contextual changes, significantly enhancing security posture and ensuring users consistently maintain the least privilege necessary.

Benefits of a Metadata Lake



IMPACT: Organizations overwhelmingly recognize the critical role of metadata lakes in reducing manual overhead, enabling automated compliance enforcement, accelerating incident response, and improving overall security effectiveness.



Current, accurate data inventory across all systems and data sets

Better data awareness for security tools

Enhanced security tool power through data sensitivity awareness

IMPACT: Organizations recognize that accurate data inventories, enhanced data context for security tools, and increased sensitivity awareness significantly improve security posture, streamline operations, and enable proactive risk management.

Stats from the Bedrock 2025 Enterprise Data Security Confidence Index

The Power of Metadata Lakes: Transforming Data Security and Governance

BEDROCK SECURITY

Implementing a metadata lake offers significant benefits in security, governance, and data management:

Security

A metadata lake enhances security by enabling faster and more efficient detection and response times across multi-cloud and large data environments, including applications in AI and cloud services. Your security team can quickly identify and address vulnerabilities by providing comprehensive visibility into data assets and their interactions. A metadata lake allows investigators to rapidly ask questions and pivot off the signals available instead of having to export metadata to Excel.such as identifying sensitive data that adversaries may have accessed or determining data at risk for unauthorized access. Enhanced security measures protect organizational assets, maintain customer trust, and ensure compliance with data protection regulations.

Governance

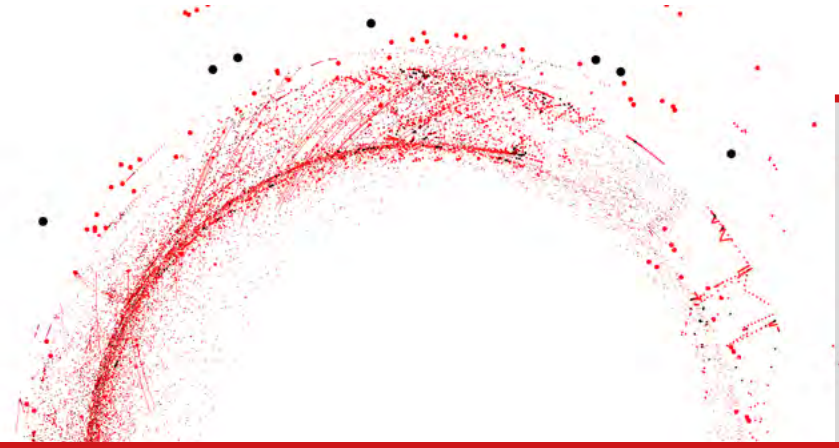
Metadata lakes promote more consistent policy monitoring across all data environments without the need for manual oversight or multiple tools. This consistency ensures that governance policies are uniformly applied, reducing the likelihood of compliance failures. You can readily identify which lines of business may fail data controls during compliance or internal audits, allowing for proactive remediation. Effective governance fosters a culture of accountability and transparency, supporting ethical data practices and enhancing organizational reputation.

Data Optimization

By providing a holistic view of data assets, metadata lakes facilitate the optimization of data use, storage, and management across your entire organization. This optimization reduces inefficient data usage and storage costs while minimizing data quality issues. You can address critical questions such as how to manage data more efficiently, leading to improved operational performance and cost savings. Optimized data management supports agile decision-making and enhances the organization's ability to innovate and compete in the marketplace.

Conclusion

A metadata lake is not just a tool—it is a strategic enabler of modern security, governance, and compliance. As your organization continues to embrace digital transformation, the ability to manage and secure data effectively across diverse environments will determine your success in mitigating risk and ensuring regulatory compliance. By adopting a metadata lake strategy, you can achieve faster security response



Next Steps

For more information on how **Bedrock Security** leverages metadata lakes for **data security and governance and management** visit **Bedrock Security**.

About Bedrock Security

Bedrock Security, the ubiquitous data security and management company, accelerates enterprises' ability to harness data as a strategic asset while minimizing risk. Its industry-first metadata lake technology and AI-driven automation enable continuous visibility into data location, sensitivity, access and usage across distributed environments. Bedrock's platform continuously catalogs data, enabling security, governance and data teams to proactively identify risks, enforce policies and optimize data usage — without disrupting operations or driving up costs. Trusted by leading financial institutions, healthcare providers and Fortune 1000 companies, Bedrock Security empowers organizations to improve data security posture management (DSPM), confidently deliver Responsible AI initiatives, and manage the exponential data growth. Learn more at www.bedrocksecurity.com