White Paper

Everpure

Take Control with AI-Ready Data

Discover, classify, and contextualize sensitive data at scale

Everpure

Contents

Modern enterprises face an uncomfortable reality in the AI era: they lack absolute visibility and granular control over their most valuable business asset—data. They need to know what data they have, where it lives, how it is related, and whether it can be safely and efficiently used by AI. As organizations launch generative AI, agentic AI, and AI data pipelines, the need for governed, relevant, secure, sovereign, and trusted data across hybrid environments becomes critical.

Regaining control with AI-ready data

AI acceleration depends on a data foundation that preserves privacy, strengthens cyber resilience, reduces data movement, and enables governed access to trusted enterprise data. This foundation fosters rapid corporate innovation without sacrificing international regulatory compliance or breaking core operational security frameworks.

The modern data visibility crisis

Three deceptively simple questions expose a critical infrastructure crisis for leadership teams:

1. **What data** do we actually have?
2. **Why** does it matter to our business?
3. **Where** is it currently located?

Answering these questions has challenged legacy discovery software. For AI, visibility alone is not enough. Enterprises must also understand data meaning, lineage, relationships, sensitivity, and business context so agents and AI applications can use the right data for the right purpose.

Decades of uncontrolled corporate data growth and shifting pipelines have produced millions of untracked, orphaned data copies scattered across disparate organizational systems. Sensitive records that originated in highly secure, isolated databases have trickled through complex extract, transform, and load (ETL) procedures, quietly embedding themselves across deep corporate file shares without the tracking or awareness of IT and information security teams.

The real-world challenge

Silos everywhere → Visibility across applications, cloud, storage, and legacy systems is fragile.

Context vacuum → Tools see data patterns but miss business meaning and relationship.

AI risk → Agents and models may use sensitive, stale, copied, or low-trust data.

Hybrid multicloud environments compound this architecture challenge. Organizations concurrently run public cloud infrastructure subject to varying regulatory updates, on-premises workloads bound by separate corporate governance rules, and legacy mainframe systems operating on completely distinct security protocols. This lack of centralized oversight fragments visibility, creating dangerous compliance liabilities.

The critical role of context in AI-ready data architecture

Legacy data discovery tools rely on basic pattern matching to identify string data types, but they lack business context. For example, simply determining that a string of characters matches a credit card number structure is not enough to assess true organizational risk. Context is what transforms raw data discovery into actionable, high-fidelity, intelligent data.

Consider the following scenario: A string of credit card numbers belonging to an EU citizen, stored unencrypted within an overseas or North American data center, represents an immediate, high-priority compliance violation that requires risk-based, automated remediation. Without business, governance, and relationship context, enterprises are forced to apply uniform, resource-intensive treatment to all discovered data, stalling analytical performance, slowing AI initiatives, increasing risk, and driving up overhead costs.

A contextual foundation is vital for enterprise AI model integrity. Understanding the precise lineage, boundaries, and characteristics of data feeding into retrieval-augmented generation (RAG) and machine learning training pipelines is critical to preventing model contamination, algorithm bias, and sensitive data leakage. Once toxic or noncompliant data enters an operational AI model, extracting it is significantly more complex and costly than preventing its ingestion at the source.

Access control imperatives

Dynamic data governance must balance real-time data accessibility for business groups with absolute security and regulatory compliance. Secure, policy-driven access decisions must continuously account for:

- **The requestor's identity and role**: Enforcing fine-grained access parameters based on organizational clearance
- **Business justification**: Verifying the explicit purpose of data utilization before provisioning access
- **Access methods and locations**: Evaluating the security posture of the endpoint, network location, and destination target
- **Continuous access validation**: Moving past point-in-time checks to real-time, active entitlement monitoring

To execute these rules, sensitivity labels and policy attributes must persist with data—files, objects, databases, schemas, and applications—so governance can follow data throughout its entire operational life cycle. When policy indicates that data should not be accessed or extracted, enforcement should be triggered through integrated controls such as data loss prevention (DLP), access control list (ACL) updates, IT service management (ITSM) workflows, security information and event management (SIEM) alerts, and downstream security tools.

However, basic labeling is merely the starting point; true resilience requires labeling to be continuously enriched with multidimensional context to drive automated data management.

The strategic data governance approach of Everpure

A mature governance strategy requires answerability across three core vectors before data is ever exposed to third-party applications or AI models.

1 Historical data management

- Where are historical, sensitive datasets stored across our legacy footprint?
- Which datasets remain exposed or accessible outside of our approved production systems?
- Where are the active pockets of shadow IT and unauthorized data duplicates that require immediate control plane intervention?

2 Risk assessment and prioritization

- What specific data estates pose the greatest reputational, operational, or regulatory risk if compromised?
- Which datasets add measurable business value, and which are purely liabilities that have passed their useful life cycle?
- How do we defensibly segment, archive, or delete nonessential storage assets to shrink our corporate attack surface?

3 Compliance and continuous governance

- How do we track what is sensitive, verify its classification status, and calculate its associated business value in near-real time?
- Which regional compliance frameworks—for example, the General Data Protection Regulation (GDPR), Payment Card Industry Data Security Standard (PCI-DSS), Health Insurance Portability and Accountability Act (HIPAA), and Digital Operational Resilience Act (DORA)—must these assets align with to remain secure?
- How do we seamlessly orchestrate automated data sharing agreements, encryption standards, tokenization policies, and comprehensive metadata cataloging across multicloud environments?

Systematic implementation methodology

Everpure™ Data Intelligence (formerly 1touch) enables a four-phase approach to AI-ready data governance: discover and classify, analyze access and risk, apply policy-driven actions, and continuously orchestrate remediation.

Everpure four-phase system method

Phase 1: Baseline assessment (at-rest and in-motion mapping)

Phase 2: Access right and overprivileged entitlement analysis

Phase 3: Attribute-based action (DLP, masking, and ACL updates)

Phase 4: Continuous orchestration and automated SIEM/ITSM alerts

Phase 1: Comprehensive sensitivity assessment

Data Intelligence establishes a real-time baseline of all data criticality, applying sensitivity tags and regulatory tracking markers across all environments. Crucially, this assessment is performed simultaneously for **data at rest** within physical or cloud storage arrays and **streaming data in motion** flowing across network fabrics. This ensures classification integrity is maintained dynamically as files replicate or shift across systems.

Phase 2: Access right analysis and control

Data Intelligence maps current user access patterns against baseline security requirements. It automatically exposes overprivileged accounts, identifies data owners, and enforces mandatory entitlement recertification. Advanced control plane algorithms calculate localized risk scores, deploy automated anomaly detection, and evaluate data loss probabilities.

Phase 3: Protective action implementation

The system converts real-time risk scores into automated protective measures. It refines data labels, enforces native DLP parameters, and deploys Attribute-Based Access Control (ABAC) policies. For example, rather than altering sensitive files directly and breaking business logic, Everpure modifies underlying ACLs to restrict access to approved groups while preserving full audit trails.

Phase 4: Orchestration and incident response

Continuous monitoring loops and automated response scripts are deployed across the enterprise data estate. When a policy violation or anomalous exfiltration attempt is detected, Data Intelligence instantly triggers preconfigured remediation workflows, isolates compromised records, and interfaces with centralized security orchestration layers to resolve data infrastructure risks before they manifest as catastrophic breaches.

Customer success: Multiplatform visibility at enterprise scale

When a global retail and fashion conglomerate found its customer data scattered across dozens of unmapped environments, it realized an uncomfortable operational reality: its most valuable resource had evolved into its biggest compliance liability. This case study details how the contextual, automated governance workflows of Everpure Data Intelligence functioned at scale to deliver real-time infrastructure security.

Hybrid multiplatform ecosystem

Cloud core: Microsoft 365 E5/M365, Purview, MIP infrastructure

Analytics layer: Snowflake cloud data warehouse systems

Legacy footprint: Distributed, on-premises storage area network/network-attached storage (SAN/NAS) file shares

The infrastructure challenge

The organization operated a highly sophisticated, multiplatform hybrid landscape. While it had made significant investments in enterprise licensing and native tools, its tools lacked the capability to discover, track, or classify sensitive personally identifiable information (PII) outside of native cloud environments. Specifically, this included customer PII within the organization's extensive data warehouse platforms and multicontinent, on-premises storage arrays.

Marketing teams required rapid, reliable access to customer purchase histories to power data-driven campaigns, but compliance teams could not verify where this data lived, who had access to it, or whether sharing it violated regional privacy regulations.

The implementation solution

Everpure Data Intelligence future-proofed the retailer's infrastructure by extending native data intelligence directly into its hybrid environment by integrating with existing labels and discovering sensitive data across on-premises file shares, data warehouses, and relational database management systems.

Business outcomes

Hours to value → Full readiness within hours

Total visibility → Direct application link

Zero dark data → Petabytes of unstructured mapping

AI guardrails → Safe RAG pipelines sans PII leak

Data Intelligence scanned multiple petabytes of unstructured format data, exposing hidden "dark data" repositories that had been overlooked for years. By achieving a validated **98.6% classification accuracy** via proprietary contextual AI and supervised machine learning models, the organization successfully mapped its entire inventory.

This allowed leadership to set granular access controls on generative AI training loops and RAG workflows, helping prevent internal models and RAG workflows from using protected consumer records outside approved policies.

Enterprise data security posture management and AI data readiness in action

Traditional, siloed security tools fail because no two enterprise data estates are identical. Every organization operates unique schema classifications, divergent access patterns, and highly complex hybrid-storage footprints. The enterprise data security posture management (DSPM) architecture of Everpure addresses this reality by shifting operations from reactive firefighting to automated risk management.

Real-world security scenario

Consider a standard operational scenario encountered by an enterprise security analyst: a targeted search across the hybrid footprint reveals critical payment card data scattered across 319 unmapped files. Further analysis via Everpure multidimensional graph capabilities exposes deep risks: select files contain personal records belonging to 43,000 unique data subjects, with over 321,000 credit card instances exposed across unauthorized, unencrypted Google Drive folders and legacy network directories.

Through the Data Intelligence advanced search plane, compliance officers construct declarative queries to verify data posture against strict international frameworks. The system instantly produces audited listings of matching data sources, filtering across diverse parameters.

Sample enterprise DSPM data inventory output

Data Source	Criticality	Data Elements Found	Regional Tag
Cloud_Array	9.5	Bank Name, CVV, Full Name	PCI-DSS/GDPR
OnPrem_NAS	8.7	CC Numbers, Account IDs	DORA Compliant

Jurisdictional intelligence and DSAR automation

A primary challenge in global regulatory compliance is differentiating between identical usernames across separate geographical regions, such as a data subject named John Smith residing in London versus a John Smith located in Geneva. By cross-referencing multidimensional contextual attributes within the storage plane, Everpure identifies distinct jurisdictional parameters automatically.

This intelligent mapping delivers two vital business capabilities:

1. **Automated DSAR fulfillment**: Compliance teams know exactly where to locate and extract every instance of an individual's data across global multicloud clusters, collapsing fulfillment times from **45 days down to single-click operations**.
2. **Breach readiness and cost control**: If an unauthorized data access event occurs, remediation can be executed within strict regional time frames. Furthermore, Data Intelligence minimizes discovery costs; during large-scale operations, it reduces data warehouse scanning costs **from $4,500 down to less than $10 per instance**, while automatically flagging and filtering fraudulent or bot-generated accounts.

Automated target relabeling policy

IF:
[Data Element = Credit Card] AND [Location = S3 / On-Prem Share]

THEN:

1. Apply persistent sensitivity label via control plane
2. Trigger real-time security alert to Splunk/SIEM ecosystem
3. Log full audit session and restrict ACL permissions via ABAC

Extending legacy tools across the enterprise with Everpure Data Intelligence

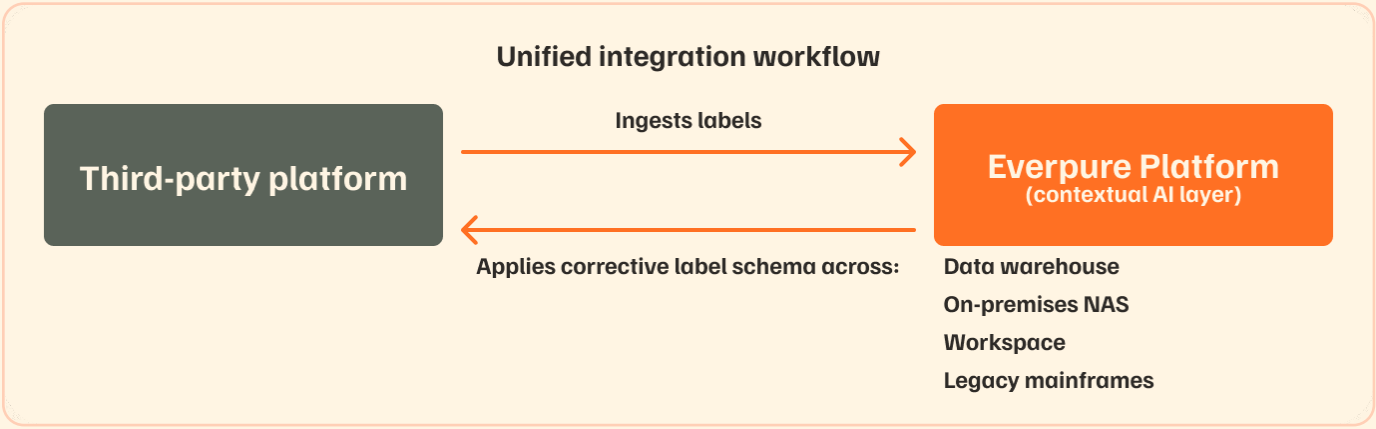
While legacy tools provide DLP capabilities within dedicated software-as-a-service and cloud structures, modern enterprises manage complex, highly fragmented landscapes that extend far beyond a single ecosystem. Critical data assets live across local on-premises servers, third-party cloud services, enterprise applications (for example, SAP, Salesforce), legacy mainframes, and disparate NoSQL databases.

This fragmentation introduces structural visibility gaps. An enterprise may achieve perfect classification compliance within its productivity cloud while remaining completely blind to unprotected customer datasets accumulating across other cloud storage, local file servers, or unstructured engineering environments.

Two-way label mapping and automated correction

Everpure Data Intelligence resolves this operational challenge through a deep, two-way integration framework with native tools. Rather than replacing your existing security software or forcing a costly configuration overhaul, Everpure ingests your established sensitivity labels (for example, Confidential, Internal Use Only, Top Secret) and applies them consistently across your entire infrastructure footprint.

Data Intelligence scans your multicloud and on-premises footprint using identical classification logic. Crucially, the system features an **automated label correction** engine. End users frequently apply incorrect sensitivity tags to files due to operational haste or not understanding corporate compliance rules. For all native file formats (including PDF, DOCX, and XLSX), Everpure evaluates the actual underlying content, automatically overrides human classification errors, and applies the correct persistent label seamlessly without interrupting workflows. This allows native tools to serve as the front-end DLP interface for end users while Everpure extends centralized control across the wider enterprise estate.



Core capabilities and technical foundation

Data Intelligence operates on a unified AI foundation built on discovery, classification, policy management, and semantic contextualization incorporating advanced large language models, neural networks, natural language processing (NLP), and supervised training mechanisms to drive data-centric orchestration:

- **Enterprise DSPM**: Continuous, agentless data posture assessments across hybrid-cloud platforms, multiregion database clusters, and legacy environments
- **Multimodal DLP**: Native DLP enforcement for both historical data at rest and streaming data in motion, adapting policies dynamically based on user identity and access context
- **Compliance automation**: Prebuilt, automated reporting pipelines that cover global regulatory standards, including GDPR, the California Consumer Privacy Act (CCPA), HIPAA, PCI-DSS, DORA, and the NIS2 Directive, while instantly automating data subject access request (DSAR) workflows
- **Extended mainframe support**: Deep, specialized structural scanning engines designed for legacy enterprise mainframes that store core transactional records
- **Third-party risk assessment**: Unrivaled visibility into how sensitive corporate assets flow across external environments, enabling secure data federation and proactive vendor risk management

Everpure centralized integration ecosystem

Security and incident ops: Splunk, ServiceNow ITSM ticket orchestration

Data catalog and analytics: Collibra, Alation operational integration

Core cloud infrastructure: IBM Db2, Amazon Redshift, Oracle, Google

Tangible business value

1 Measurable attack surface reduction

By scanning historical records and tracking deep context, organizations gain absolute visibility into data age, usage metrics, and business relevance. Security teams can confidently identify, archive, or permanently delete multi-terabyte datasets that have passed their useful life cycle, drastically cutting storage infrastructure costs and shrinking the corporate risk profile.

2 Maximized ROI on existing toolsets

Everpure builds on your existing investments, maximizing the capability and accuracy of deployment frameworks without requiring a costly rip-and-replace strategy. It refines your data pipelines with actionable telemetry, allowing security tools to operate more efficiently.

3 Enterprise-grade AI readiness

Proper data classification is the fundamental cornerstone of safe corporate AI adoption. By ensuring that structured and unstructured data landscapes are continuously mapped, labeled, and monitored at the core storage layer, Everpure gives your enterprise the absolute confidence to deploy next-generation generative AI technologies, leverage full data value, and secure your digital future.

Regulatory compliance quick reference

Data Intelligence provides direct contribution matching across all major global regulatory frameworks.

Regulation framework	Operational target	Everpure Data Intelligence capabilities
ISO/IEC 27001	Information security management	Surfaces and classifies hidden data assets; supports unified access control and audit readiness
NIS2/DORA	Information and communication technology (ICT) resilience and supply chain	Tracks sensitive data across hybrid clouds; monitors third-party data movement and unauthorized data sharing
GDPR/CCPA	Personal data protection	Locates PII; supports automated data minimization, breach readiness, and single-click DSAR fulfillment
HIPAA/Art. 9 GDPR	Protected health information (PHI)	Automatically tags and monitors PHI data fields; enforces consent parameters and compliant audit trails
PCI-DSS 4.0	Payment card security	Identifies cardholder data; reduces audit scope via automated encryption and real-time access restriction

Take control of your data destiny

Everpure helps enterprises turn fragmented, sensitive, and under-governed data into AI-ready, intelligent data, safeguarding the hybrid records of over 500 million individuals globally. Through deep, intelligent data capabilities, broad ecosystem integration, and strategic partnerships, Everpure empowers modern enterprises to convert data risks into their strongest long-term competitive advantages.

The blueprint is clear: enhance your existing infrastructure investments to achieve comprehensive data governance, or continue struggling with fragmented visibility and reactive policies. The choice defines your organization's data security and AI readiness posture for years to come.

Learn More About Everpure Data Intelligence

Visit Our Website | **800.379.PURE**

ID26104-01-en-06/26