SOLUTION PROFILE

Simplify Data Privacy and Compliance with Automated Discovery & Protection

Rapid Data Discovery for Privacy and Compliance Lowering Risk and Cost

Sensitive Data Discovery

Find sensitive data using rules based on business context and privacy policies.

Eliminate Duplication

Deduplicate data copies helping reduce the surface area for breaches while lowering storage cost.

Open Ecosystem

Integrate with Pentaho™ and third party specialty tools to help with redaction, pseudonymization and movement.

Protect all Personally Identifiable Information (PII), Personal Health Information (PHI) and other sensitive data for privacy and regulatory compliance (GDPR, CCPA etc.) across your organization.

Hitachi Vantara's Pentaho platform is a powerful data management platform that automates large-scale data discovery and classification to manage across structured and unstructured data sources. This allows organizations to apply privacy and compliance rules to sensitive data sets natively, or via API connectivity to governance tools. The platform finds potentially sensitive information, applying rules to data based on business context and your security and privacy policies. By using machine-learning (ML) to apply AI fingerprinting Pentaho can quickly identify potential data duplication speeding up the process and reducing the risks to your organization. The platform is designed to be deployed on-premises or in the cloud while integrating with Pentaho as well as third party specialty tools to help redaction and data movement tools.

Lower Costs

AI-driven automatic identification and enforcement at massive scale (Petabytes, hundreds of millions of columns)

Lower Risks

Meet privacy and compliance requirements, applying business specific data rules that fit your needs

Higher Business Value

Find fit-for-purpose data quickly for faster time-to-value

Symptoms? You May Have a Problem:

As the distribution of data grows many organizations struggle with easily knowing what is sensitive in their structured and unstructured data repositories. Large volumes of duplicate data that can sometimes harbor unknown and unmanaged sensitive data leaving enterprise at risk. Moreover, organizations need help with what copies to retain and what information to redact. Identifying sensitive data can require large amounts of manual efforts, making it costly and delaying the time it takes to identify and classify sensitive information. Modernization efforts become more expensive as data and security team grapple to know what data to move to the cloud. Compliance regulations must always be met to avoid unnecessary business impact. This can lead to complex discovery requests involving exhaustive effort to classify structured and unstructured data.

Challenges Addressed:

- Accelerate sensitive information identification across all data objects with limited manual efforts
- Examine & scan all data, structured, unstructured, and semi-structured across on-premises, Hadoop and cloud environments
- Search a catalog of accurately tagged data according to lifecycle and security policies
- Reduce risk by eliminating duplicate data containing sensitive information and by ensuring cloud modernization projects align to security policies
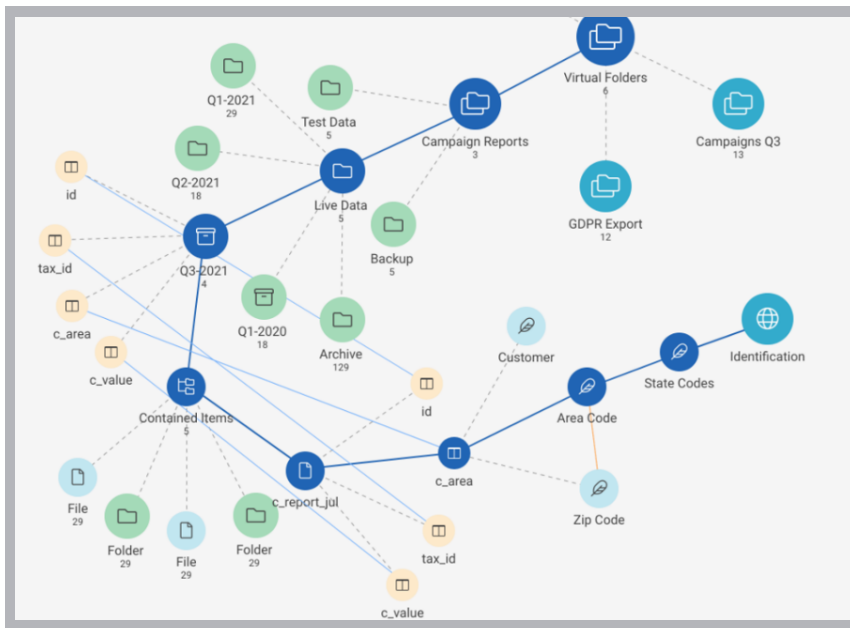- Reduce costs and simplify compliance through localized data processing

Figure 1: Automate content discovery and classification at scale

Figure 2: Data Integration- into Pentaho and 3rd party specialty tools to help with redaction and data movement.

Proven Platform to Discover feature to meet privacy and compliance include:

- Business glossaries representing your organization's data for automated assets discovery and user-friendly search

- Programmable and machine learning driven discovery uncovering and exposing any potential sensitive data

- Apply business rules fitting your regulatory needs and corporate privacy policy rules actioned and applied across the complete data estate

- Integration with existing third-party compliance and privacy tools

- Automated data profiling surfacing data duplication and redundancy for better storage and retention management

- Leverage quickly assessed data quality and lineage providing trusted data self-service across your organization

- Automation providing affordable flexibility and scalability where demand is unpredictable and labor intensive

- A single source of truth for governance teams and data developers to find data fit-for-purpose

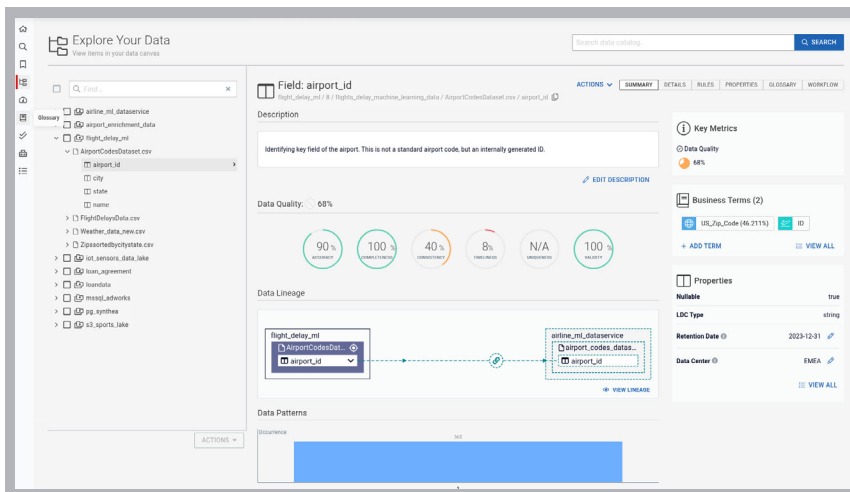- API-first architecture providing metadata to drive all your dynamic data management and governance processes

Contact Us →

Learn more about how Pentaho can help you meet you Privacy and Compliance needs at scale while reducing the cost to do so. Meet with an expert today to see the full solution functionality including discovery, classification, quality, sensitive data protection and more.

ABOUT HITACHI VANTARA

Hitachi Vantara, a wholly-owned subsidiary of Hitachi Ltd., delivers the intelligent data platforms, infrastructure systems, and digital expertise that supports more than 80% of the Fortune 100. To learn how Hitachi Vantara turns businesses from data-rich to data-driven through agile digital processes, products, and experiences, visit hitachivantara.com.

Hitachi Vantara
A Hitachi Group Company