

### Manage Data, Achieve Compliance, & Accelerate Productivity

#### Discover and Explore

- Easily track, classify, and locate data within the unified metadata repository

#### Optimize

- Get the best results from Hadoop with instant workload insights and optimization guidance

#### Audit

- Maintain a full audit history across all of Hadoop in a single place

#### Trace

- Visualize the upstream and downstream lineage of data to verify reliability

#### Protect

- Keep sensitive data secure with performance-optimized encryption and key management

#### Manage Lifecycle

- Define and automate complex data lifecycle activities with integrated metadata policies

“We are the custodians of financial information for our customers when they’re trusting us to transfer money to their near and dear ones. We need to make sure that we are compliant and have proper monitoring and auditing, and that’s where [Cloudera] Navigator comes in.”

Western Union

## Cloudera Navigator

### Integrated Data Management and Governance for Hadoop

Big data represents a major opportunity for enterprises to become information-driven in their strategies, operations, and customer engagements - with the ability to store and analyze all types of data at any scale and open it up to more users and analytic tools. However, this opportunity can also lead to data management challenges, with armies of business users wanting self-service access to discovery data; administrators needing to know how data is being used to optimize analytic performance; and security teams requiring visibility into how data is accessed and used to meet compliance regulations. An enterprise big data platform must be able to address the data management and compliance needs across the organization, without limiting the flexibility and benefits of big data itself.

Cloudera Navigator is the only integrated data management and governance solution for big data and Apache Hadoop. Built into the core of Cloudera Enterprise, Cloudera Navigator is trusted by hundreds of users to get unprecedented visibility into their data and provides the auditing and data protection necessary to comply with even the most stringent regulatory restrictions. Powered by the only comprehensive metadata foundation for Hadoop, it automatically brings together all the technical metadata from each platform component and the user-defined business metadata across the organization into a single, searchable repository. From this, Cloudera Navigator provides four fundamental components for effective data management:

### Self-Service Data Discovery & Analytics

Business users can effortlessly find and trust the data that matters most

- Discover and explore data across the only unified metadata repository with intuitive full-text search and SQL access
- Locate data sets based on business context and classification, combined with automatic technical context - making it easy to find similar, relevant data

### Active Data Optimization

Database Administrators can get instant insights to optimize the most critical workloads

- Instantly analyzes existing SQL logs for comprehensive visibility into which queries are the most critical, what data is accessed the most often, and how is that data used.
- Improve performance and efficiency of Hadoop with intelligent optimization guidance
- Reduce workload development time with compatibility identification for fast success with Hadoop

### Compliance-Ready Governance & Protection

Security teams can track, understand, and protect access to sensitive data

- Automatically maintain a full audit history and track every access attempt, right down to the user ID, IP address, and full query text
- Track how data is used and changing with column-level, visual lineage to quickly identify the origin of a data set and its impact on downstream artifact
- Protect all data with high-performance encryption and key management through Navigator Encrypt and Navigator Key Trustee (for more information, download the “[Encryption and Key Management Datasheet](#)”)
- Integrate with the leading enterprise metadata, lineage, and SIEM applications, out-of-the-box

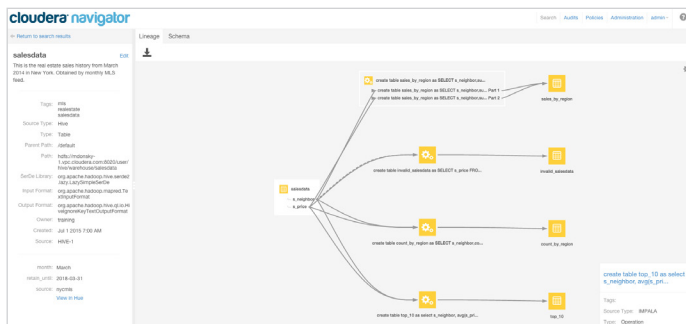
## Hadoop-Scale Data Lifecycle Automation

Data stewards can efficiently manage and enforce critical lifecycle policies, risk-free

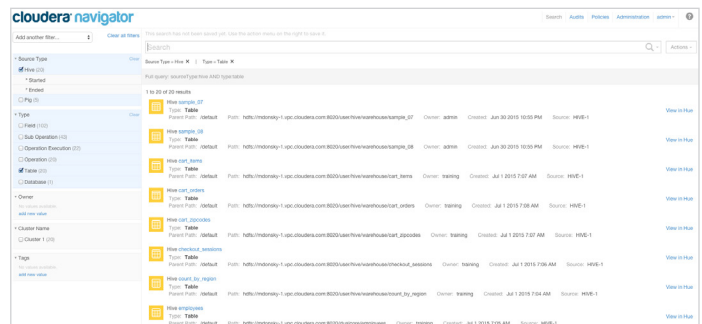
- Automate crucial stewardship and curation activities – such as metadata classification, data archiving and retention, or even invoking partner products for additional data preparation and transformation – with the flexible policy engine
- Ensure business continuity with the only built-in backup and disaster recovery
- Manage the data lifecycle beyond just Hadoop with seamless partner tool integration

## Key Features of Cloudera Navigator

Feature	Function	Benefit
<b>Discovery &amp; Exploration</b>		
Unified Technical Metadata	Bring in all metadata from HCatalog, HDFS, Sqoop, etc. into a single, searchable interface	Stay in control of growing volumes of data for consistent, precise retrieval
Comprehensive Business Metadata	Augment files, tables, and individual columns with custom business context, tags, and key/value pairs	Unlocking the value of data by classifying it in meaningful ways for various stakeholders, including business analysts, data scientists, and data stewards
Metadata Search	Search across all business and technical metadata to find assets (ie data older than 7 years, tables created by certain users, files that contain sensitive data)	Find diverse data through a common interface
<b>Audit &amp; Access Control</b>		
Audit Configuration	Automatically configure audit tracking for HDFS, Impala, Hive, HBase, and Sentry; set thresholds	In a few simple clicks, ensure all necessary audit data is captured
Audit Dashboard	Visualize and summarize data access via a simple, queryable interface	Single place to quickly identify outliers and security breaches
Third-Party Integration	Integrate audit information for use in global Security Information and Event Management (SIEM) systems	Integrate data from Hadoop into infrastructure-wide reporting and processes
<b>Lineage</b>		
Intuitive Visualization	View upstream and downstream lineage in an easy-to-follow graph at the file, table, and individual column-level	Quickly identify the origin of a data set and the impact on downstream operations
<b>Encrypt &amp; Key Trustee</b>		
Transparent Encryption	Protect all data without the performance impact	Open up even the most sensitive data to analytics, risk-free
Enterprise Key Management	Manage encryption keys and security assets through a separate, centralized solution	Integrate with existing HSMs and key management policies for compliance
<b>Optimizer</b>		
Active Data Optimization	Analyze workloads across all data management systems for peak performance optimizations and workload rationalization	Put the right workload in the right place at the right time for the best results with Hadoop
<b>Metadata Policies</b>		
Metadata Policy Management	Trigger actions (such as autoclassification of metadata) for specific datasets based on arrival or scheduled intervals	Easily set, monitor, and enforce data management policies and integrate with third-party tools



View upstream and downstream lineage



Search metadata for specific items of interest