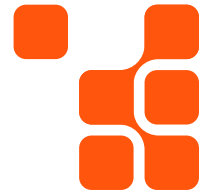


Cloudera Octopai Data Lineage: Metadata Management & Data Lineage Solution

Enterprise-Grade Visibility for Hybrid Data Ecosystems



Cloudera Octopai Data Lineage is the leading automated data lineage, discovery, and catalog solution that enables enterprise data teams to find, understand, and manage their metadata with unprecedented efficiency. Cloudera Octopai Data Lineage provides organizations with complete visibility across complex hybrid data environments, helping teams execute changes with confidence, ensure trustworthy data, and accelerate strategic initiatives. With a 100% automated approach and integration with 60+ technologies out of the box, Cloudera Octopai Data Lineage delivers value in hours, not weeks, making it the premier solution for modern data governance challenges.

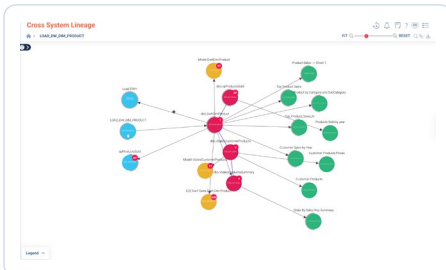
60+

Integration with 60+ technologies out of the box

Multidimensional Lineage: See Your Entire Data Journey

Cloudera Octopai Data Lineage revolutionizes metadata management through its comprehensive, multidimensional data lineage capabilities that provide full visibility into data flows at every level:

Cross-System Lineage



End-to-End Visibility

Trace data from source systems to final reports across your entire data landscape

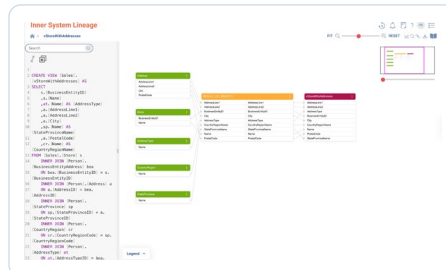
Impact Analysis

Understand upstream and downstream dependencies before making changes

Technology Flow Mapping

Visualize how data moves through different platforms and tools

Inner-System Lineage



Column-Level Detail

See field-level transformations within processes and semantic layers

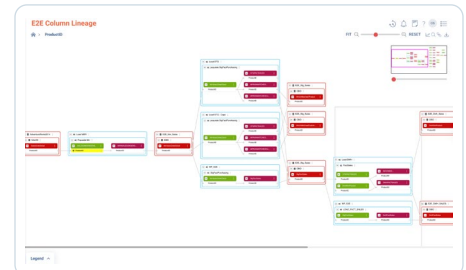
Script Visualization

Understand source-to-target transformations within SQL, ETL packages, and BI tools

Transformation Transparency

Expose all expressions and logic applied to your data

End-to-End Lineage



Asset Level Lineage

End-to-end asset level lineage between systems from the source system into the end consumer

Business Challenges Solved by Cloudera Octopai Data Lineage

CHALLENGE	BUSINESS IMPACT	SOLUTION
Cloud Migration	Costly, risky migrations with 42% failing due to hidden dependencies	Map of cross-system dependencies automatically to cut migration time by up to 75%
Business Disruption	Data errors causing decision delays and \$300K/hour in downtime costs	Reduce root cause analysis from days to minutes with visual lineage tracing
Delivery Bottlenecks	14+ hours/week lost waiting for subject matter experts	Break tribal knowledge silos with automated documentation and self-service discovery
Compliance Penalties	Manual processes extending audit prep by 72%	Turn metadata into audit-ready proof with automated lineage and governance
Unexpected Outages	Changes triggering cascading failures across systems	Predict impacts before deployment with automated dependency mapping
Cloud Spend	37% of costs wasted on orphaned or redundant resources	Identify unused tables and redundant pipelines to optimize cloud expenses



Cloudera, Inc. | 5470 Great America Pkwy, Santa Clara, CA 95054 USA | cloudera.com

Cloudera is the only cloud anywhere platform for data and AI. With 100x more data under management than other cloud-only vendors, Cloudera empowers global enterprises to transform data of all types, on any public or private cloud, into valuable, trusted insights. Our open data lakehouse delivers scalable and secure data management with portable cloud-native analytics, enabling customers to bring GenAI models to their data while maintaining privacy and ensuring responsible, reliable AI deployments. The world's largest brands in financial services, insurance, media, manufacturing, and government rely on Cloudera to be able to use their data to solve the impossible—today and in the future.

To learn more, visit [Cloudera.com](https://cloudera.com) and follow us on [LinkedIn](#) and [X](#).