Cloudera Octopai Data Lineage 1.0.0

Discovery

Date published: 2025-10-09
Date modified: 2025-10-20

CLOUDERA

https://docs.cloudera.com/

Legal Notice

© Cloudera Inc. 2026. All rights reserved.

The documentation is and contains Cloudera proprietary information protected by copyright and other intellectual property rights. No license under copyright or any other intellectual property right is granted herein.

Unless otherwise noted, scripts and sample code are licensed under the Apache License, Version 2.0.

Copyright information for Cloudera software may be found within the documentation accompanying each component in a particular release.

Cloudera software includes software from various open source or other third party projects, and may be released under the Apache Software License 2.0 (“ASLv2”), the Affero General Public License version 3 (AGPLv3), or other license terms. Other software included may be released under the terms of alternative open source licenses. Please review the license and notice files accompanying the software for additional licensing information.

Please visit the Cloudera software product page for more information on Cloudera software. For more information on Cloudera support services, please visit either the Support or Sales page. Feel free to contact us directly to discuss your specific needs.

Cloudera reserves the right to change any products at any time, and without notice. Cloudera assumes no responsibility nor liability arising from the use of products, except as expressly agreed to in writing by Cloudera.

Cloudera, Cloudera Altus, HUE, Impala, Cloudera Impala, and other Cloudera marks are registered or unregistered trademarks in the United States and other countries. All other trademarks are the property of their respective owners.

Disclaimer: EXCEPT AS EXPRESSLY PROVIDED IN A WRITTEN AGREEMENT WITH CLOUDERA, CLOUDERA DOES NOT MAKE NOR GIVE ANY REPRESENTATION, WARRANTY, NOR COVENANT OF ANY KIND, WHETHER EXPRESS OR IMPLIED, IN CONNECTION WITH CLOUDERA TECHNOLOGY OR RELATED SUPPORT PROVIDED IN CONNECTION THEREWITH. CLOUDERA DOES NOT WARRANT THAT CLOUDERA PRODUCTS NOR SOFTWARE WILL OPERATE UNINTERRUPTED NOR THAT IT WILL BE FREE FROM DEFECTS NOR ERRORS, THAT IT WILL PROTECT YOUR DATA FROM LOSS, CORRUPTION NOR UNAVAILABILITY, NOR THAT IT WILL MEET ALL OF CUSTOMER’S BUSINESS REQUIREMENTS. WITHOUT LIMITING THE FOREGOING, AND TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW, CLOUDERA EXPRESSLY DISCLAIMS ANY AND ALL IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO IMPLIED WARRANTIES OF MERCHANTABILITY, QUALITY, NON-INFRINGEMENT, TITLE, AND FITNESS FOR A PARTICULAR PURPOSE AND ANY REPRESENTATION, WARRANTY, OR COVENANT BASED ON COURSE OF DEALING OR USAGE IN TRADE.

Contents

Automated Discovery Space - Your Search Engine for Cross-System Data Asset Discovery..... 4

Export items from Discovery..... 6

Automated Discovery Space - Your Search Engine for Cross-System Data Asset Discovery

Discovery Space in Cloudera Data Lineage (formerly Octopai) is a high-performance search engine designed for data experts who need instant access to metadata, scripts, and stored procedures across multiple systems. Think of it as Goo...

Find the Data You Need—Instantly and Accurately

Discovery Space in Cloudera Data Lineage is a high-performance search engine designed for data experts who need instant access to metadata, scripts, and stored procedures across multiple systems. Think of it as Google Search for your data ecosystem, enabling fast, context-aware queries that preserve the native structure, language, and formats of the harvested systems.

Key Capabilities of Discovery Space:

1. **Comprehensive Metadata Indexing:**
    - Discovery Space represents the full scope of harvested metadata, including:
        - Tables, columns, reports, and dashboards
        - ETL processes, scripts, stored procedures
        - Business logic and dependencies
    - The search engine operates across databases, BI tools, ETL platforms, and cloud environments, ensuring full visibility into your data landscape.
2. **Intelligent Search for Precision:**
    - Supports fuzzy and exact search parameters to narrow results efficiently, allowing users to locate relevant objects without manual filtering.
    - Queries are performed using an intelligent search engine that understands system-specific syntax and maintains the original context of the metadata.
3. **Native Format & Terminology Preservation:**
    - Unlike generic search tools, Discovery Space preserves the structure and terminology of each source system, ensuring that results align with the way data is defined and stored in its native environment.
    - This is critical for impact analysis, root cause investigations, and governance workflows, where context and accuracy are essential.
4. **Downloadable Insights for Impact Analysis:**
    - Discovery Space does not alter or annotate data, but instead enables users to download search results for further scoping and impact assessment.
    - These insights can be cross-referenced in Knowledge Hub, where additional documentation, collaboration, and governance processes can take place.

Optimized for Data Engineers, Analysts, and Governance Teams

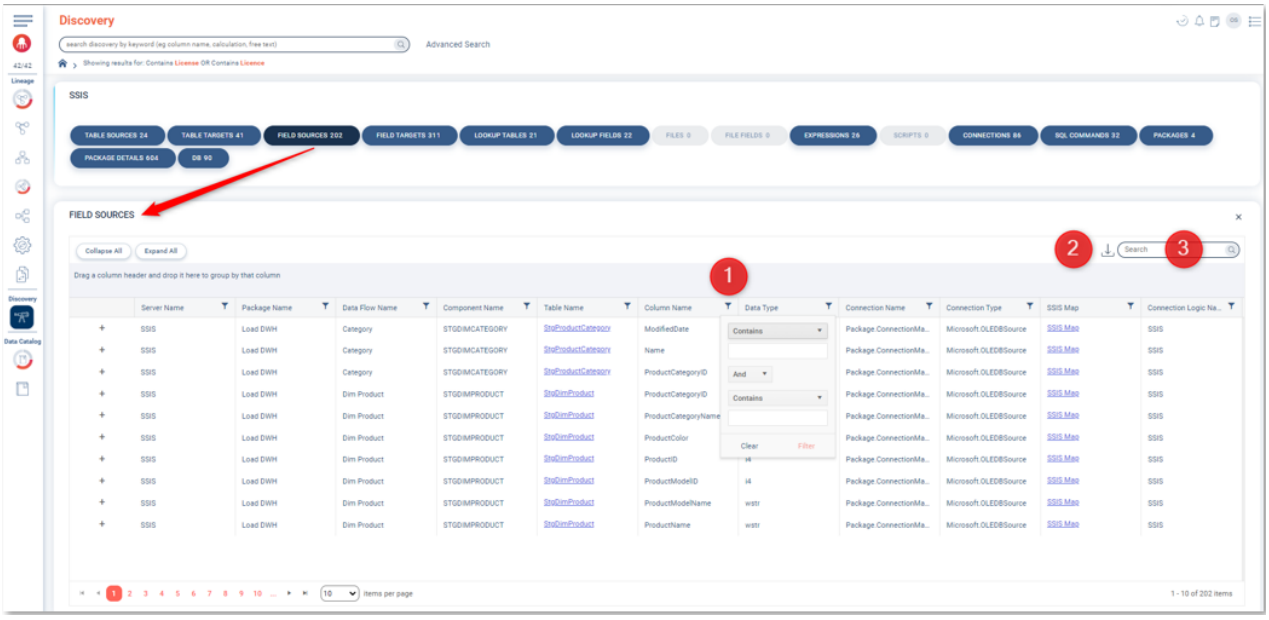
With **Discovery Space** , data professionals can: **Locate critical data assets instantly** across complex, hybrid environments **Analyze metadata dependencies** to assess risks and change impacts **Streamline investigations** with precise, system-specific searches

Systems Section:



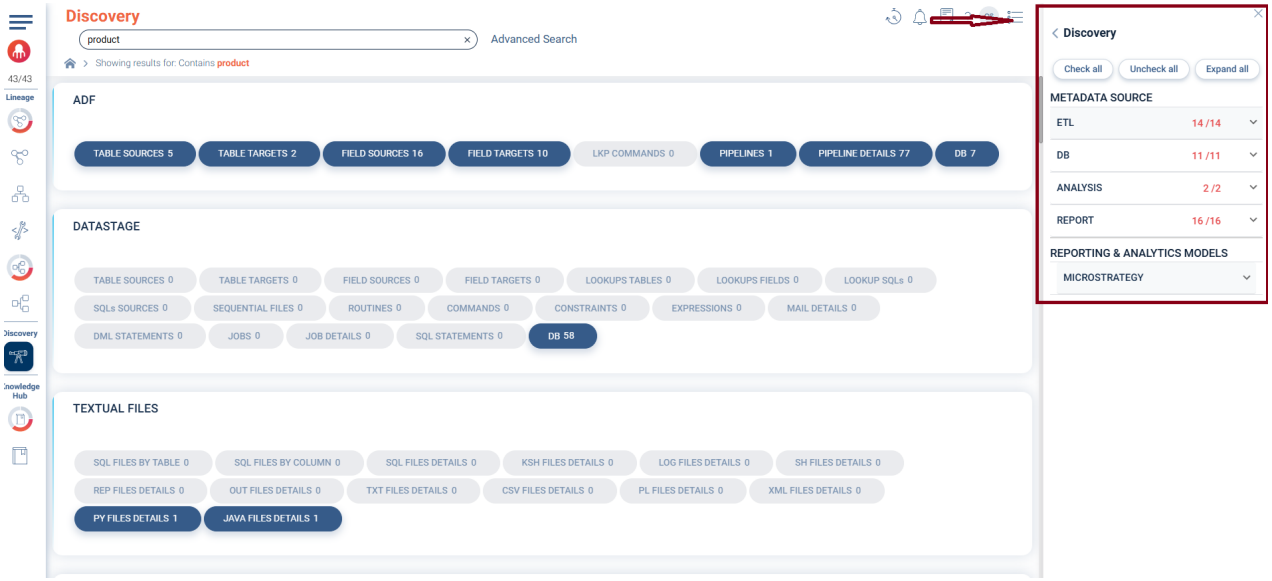
1. **Search Box** Searches for any value across the BI landscape
2. **Advance Search** Search values using AND / OR when it comes to a non-conclusive term or searching for more than one value
3. **How Cloudera Data Lineage Searches the Value** Hover over the tag to know how Cloudera Data Lineage searches the value

Detailed Level:



1. **Column Filter Funnel** Narrow down the search within a column
2. **Export the list to Excel** Use the information for your own purposes like pivot tables, workload & effort planning, etc.
3. **Specific tool search** Helps to search for a specific term within the open tool

Apply Filters on which Metadata Sources you wish to apply the discovery function:



Export items from Discovery

Learn about exporting the Discovery Module results from Cloudera Data Lineage (formerly Octopai).

Procedure

1. Search for your item in the **Discovery** module.
2. Click Download to export the list to an Excel file in XLSX format.

Figure 1: Export items from Discovery

