

Informatica Enterprise Data Catalog Advanced Scanners

Benefits

- Automatically extract the most granular metadata from a wide array of complex enterprise systems, including applications that are considered “black boxes”
- Obtain comprehensive and detailed data lineage to understand the provenance of all your enterprise data and the impact of any changes
- Access and trace data transformations throughout the enterprise to support your data-driven digital transformation initiatives

Metadata Scanners for Complex Enterprise Systems

The enterprise data landscape is highly complex, with petabyte-scale data residing across hundreds of data sources—including databases, data warehouses, legacy mainframe systems, enterprise and SaaS applications, Hadoop and Spark clusters, and data lakes—across a hybrid and multi-cloud environment. For most enterprises, this data is often trapped and siloed. It is difficult to get to and even harder to understand.

The inability to scan all data sources and extract, understand, and track data dependencies across data sources and scripting languages is a major hurdle that enterprises face today on their digital transformation journeys. Moreover, regulatory compliance, enterprise data governance, as well as a myriad of other data-driven digital transformation initiatives—including modernizing data warehouses and data lakes in the cloud—require you to have comprehensive audit trails to mitigate all types of risks. This means you need the ability to trace data and its transformation throughout the data lifecycle from source to target and across data pipelines at a granular level.

Powered by the metadata-driven intelligence in the Informatica® CLAIRE® AI-engine, Informatica Enterprise Data Catalog offers the most comprehensive set of advanced scanners designed to extract deep metadata and lineage from a myriad of hybrid and multi-cloud data sources. With end-to-end data lineage and impact analysis capabilities, you can easily analyze the data dependencies from source to target, understand the impact of proposed changes, and perform root-cause analysis of data issues.

Key Features

Rapid Data Cataloging with Deep Metadata Extraction and End-to-End Data Lineage

Informatica Enterprise Data Catalog Advanced Scanners allow you to extract metadata and derive detailed data lineage from many data sources, including the most complex systems—such as legacy and mainframe systems, SQL dialects, and various enterprise applications—across hybrid and multi-cloud environments. For instance, dynamic SQL generation, parametrized procedures, ETL code controlling data transformations, file copying scripts, and many others are typical examples where lineage usually cannot be extracted or is not detailed enough to support various data-driven digital transformation initiatives. Enterprise Data Catalog Advanced Scanners extract detailed data lineage from these complex systems and tackle the most stringent data lineage requirements.

With Enterprise Data Catalog Advanced Scanners, you can visually inspect every script, procedure, or process to fully understand its logic and internal data flow. You can obtain a complete column-level data lineage, including a full inventory of all the potential lineage sources with rich details. This enables you to understand every transformation that has occurred to the data within the enterprise. The scanners allow you to scan both static and dynamic code as well as perform language parsing to obtain automated data lineage. With this level of deep metadata insights, you can speed your data-driven digital transformation initiatives—including enterprise data governance, advanced analytics, regulatory compliance, and modernizing data warehouses and data lakes in the cloud—with full transparency, control, auditability, and confidence.

Databases and Scripting Languages

Enterprise Data Catalog Advanced Scanners enable detailed data lineage extraction from a variety of SQL dialects at scale, including stored procedures for Oracle, SQL Server, IBM Netezza, MySQL, and Teradata, to name a few. Extracted data lineage provides full visibility into the procedure calls with parameter tracking, dynamic SQL generation from values based on parameters, database queries, and more. Supported objects include views, procedures, functions, triggers, macros, external tables, and so on. Analysis can be performed online on a live database or offline using metadata-only extracts.

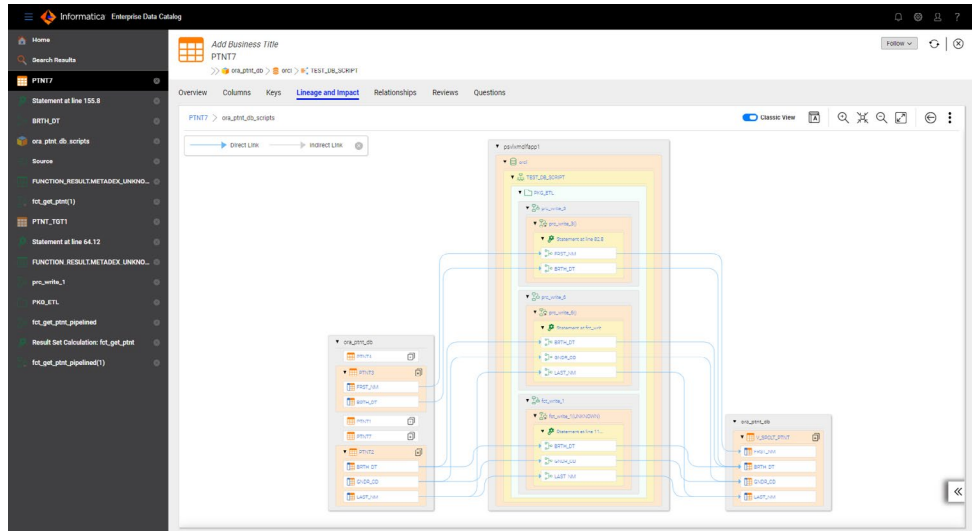


Figure 1: Deep metadata extraction and end-to-end data lineage.

ETL Tools

Enterprise Data Catalog Advanced Scanners allow you to extract metadata and lineage information from a wide range of multi-vendor ETL tools beyond Informatica’s data integration products. You can obtain end-to-end data lineage with granular-level detail including column-level lineage and transformation logic that can be seen from and to systems connected through the ETL tools. For instance, the scanner for IBM InfoSphere DataStage provides detailed data lineage supporting tasks such as runtime column propagation, local and shared containers, support for both server jobs and parallel jobs, and embedded SQL of all types and forms. You can easily visualize every component in an ETL job and gain quick access to every expression being applied to the data from source to target.

Legacy and Mainframe Systems

Custom applications built on scripts and codes, as well as legacy and mainframe systems built on COBOL, are often considered to be “black boxes.” Only skilled developers are able to understand the transformations that are applied to the data. Enterprise Data Catalog Advanced Scanners deliver the most complete, end-to-end data lineage views for your legacy and mainframe assets, enabling you to virtually eliminate black boxes. With Enterprise Data Catalog Advanced Scanners, you can extract transformation logic and provide deep insights into the lineage involved, including support for standard mainframe utilities for end-to-end data flow analysis. This enables you to open up these applications to more users for manageability, transparency, and effective governance.

About Informatica

Digital transformation changes expectations: better service, faster delivery, with less cost. Businesses must transform to stay relevant and data holds the answers.

As the world's leader in Enterprise Cloud Data Management, we're prepared to help you intelligently lead—in any sector, category, or niche. Informatica provides you with the foresight to become more agile, realize new growth opportunities, or create new inventions. With 100% focus on everything data, we offer the versatility needed to succeed.

We invite you to explore all that Informatica has to offer—and unleash the power of data to drive your next intelligent disruption.

Complex Enterprise Applications

Enterprise applications such as SAP BW, and SAP BW/4HANA are complex systems that often do not provide easily shareable descriptions of internal storage, processes, and relationships. With Enterprise Data Catalog Advanced Scanners, you can extract metadata, lineage, and relationships, making data easily accessible and understandable for all users, including data engineers, data analysts, and data stewards. You can trace even the most complex data lineage end-to-end across SAP and third-party enterprise applications, databases, and business intelligence tools, regardless of whether they are on-premises or in the cloud. You can gain comprehensive visibility into how data is transformed across SAP technologies as well as in linked systems.

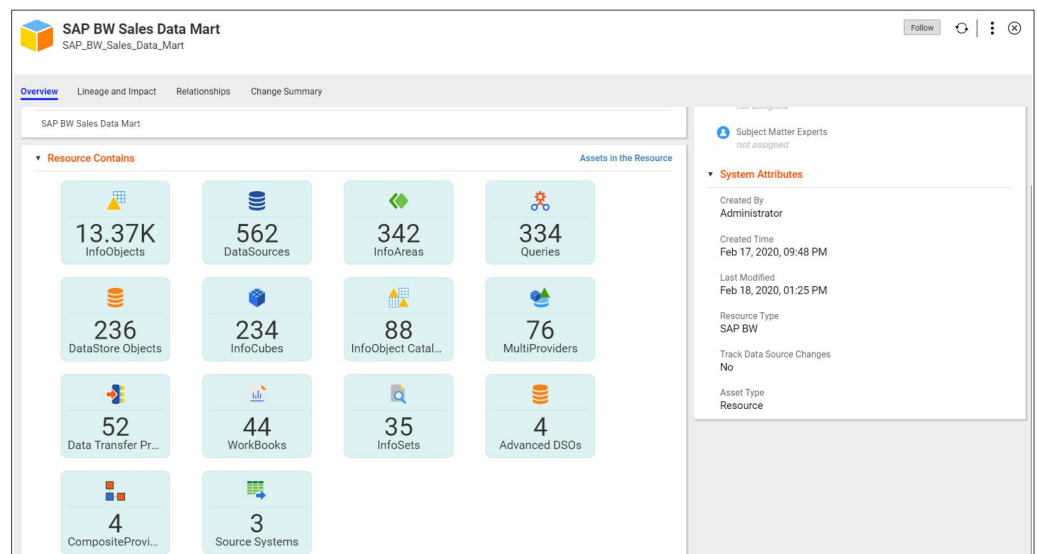


Figure 2: Complete SAP object and relationship ingestion.

Next Steps

For more information, visit the [Informatica Enterprise Data Catalog Advanced Scanners](#) web page.



Worldwide Headquarters 2100 Seaport Blvd., Redwood City, CA 94063, USA Phone: 650.385.5000, Toll-free in the US: 1.800.653.3871

IN06_0220_03844