

Informatica Data Explorer

Understanding the True Nature of Your Data Sources

Before data can be integrated or used in a data warehouse, CRM, ERP or business analytics application, its content, quality and structure must be understood. Too often, mission-critical IT, data integration and master data management initiatives experience overruns or outright failures because they are based upon an inaccurate or incomplete understanding of the source data. This happens because databases are complex, the volumes of metadata and data can be vast and difficult to unravel, and the process of understanding source data can be laborious and prone to error.

BENEFITS

- Delivers accurate source system knowledge
- Improves corporate data quality and accuracy
- Enables accurate data migration, integration, and consolidation
- Helps integrate multiple, disparate data sources
- Accelerates IT project completion time
- Deploys quickly and generates results via wizards
- Reduces risk in data-driven projects and enhances IT investments and ROI
- Provides key information about sources to drive data cleansing with Informatica Data Quality

Informatica eliminates this age-old problem with its market leading data profiling suite **Informatica Data Explorer**. Informatica Data Explorer puts powerful data profiling and mapping capabilities in the hands of the business and IT. Data analysts and data stewards work with Informatica Data Explorer to search and profile data from any source. With unparalleled ease of use, the software reveals hidden data quality issues, gaps, and inconsistencies and incompatibilities within data sources. Accurately profiled data accelerates and streamlines data integration initiatives and other business or IT projects. Organizations use Informatica Data Explorer to create a complete and completely accurate picture of the content, quality, and structure of their enterprise data that can be used as the foundation for addressing data quality and data integration enterprise-wide.

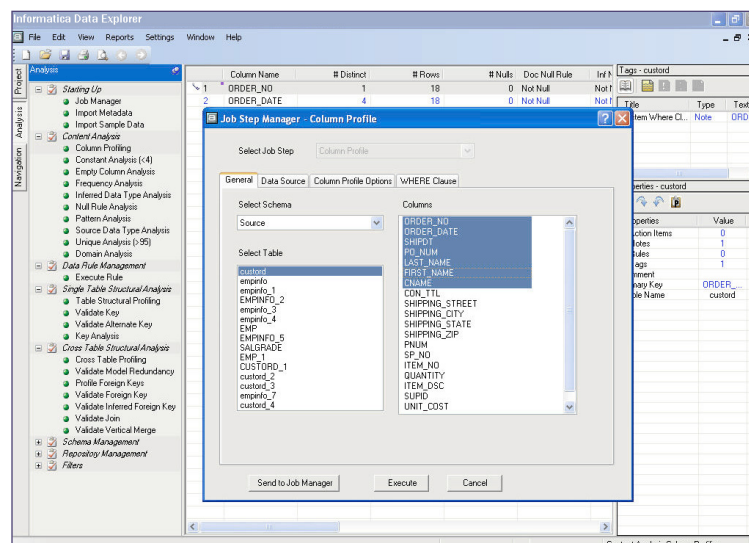


Figure 1: Informatica Data Explorer

Data Profiling Overview

Data profiling with Informatica Data Explorer is the essential first step in any business initiative leveraging data and technology. By quickly providing an accurate profile and mapping specifications for all data sources, Informatica Data Explorer dramatically accelerates data integration lifecycles, de-risks the implementation of new applications and databases, and provides a starting point for improving data quality. The solution has been successfully deployed in hundreds of data integration, data migration, data warehouse and data quality discovery projects world-wide.

Put simply, Informatica Data Explorer enables organizations to discover the content, quality and structure of all their corporate data. The software identifies existing data anomalies and creates a normalized schema capable of supporting your data. The solution will also generate data cleansing and data transformation specifications that can be deployed using Informatica Data Quality and Informatica PowerCenter.

Data profiling with Informatica Data Explorer helps organizations to answer important questions:

- Do we have the data necessary to complete the project on time and on budget?
- Does the data definition support our business requirements?
- Will the project be able to cost-effectively produce and maintain the information required by the business?
- Does the data consistently and accurately represent the business needs?
- Will the relationship between the data elements support the business requirements?
- Will we be able to integrate, consolidate, aggregate, cross-reference and pivot the data for usable reports?
- What data needs to be cleansed?
- What data needs to be transformed?
- Will the data be correct, consistent and stable?

Ease of Use

Although Informatica Data Explorer is revolutionary, its data profiling and mapping methodologies are straightforward. The goal of the profiling process is to provide accurate metadata and a complete understanding of the data as it actually exists — not how it was designed years or decades ago. Since its design, the data will certainly have been altered and the existing documentation may no longer accurately describe the content and structure of the data source. Following the steps of the Informatica data investigation process, the data analyst can discover the true content, quality and structural relationships of the source and how it best maps to an accurate data model or to a predefined target (such as a packaged CRM system, ERP package or data warehouse design).

Informatica Data Explorer is quick to deploy and quickly generates results, as data profiling processes are run through wizards that bring users through all the possible options in a logical and easy to understand workflow. Users can create processes and steps and save these to be run as interactive jobs or as scheduled batch processes.

Broad Profiling Capabilities

Informatica Data Explorer analyzes data in three dimensions - down columns, across rows and between tables - to discover the true content, structure and quality of the data and produce a fully normalized model of the source data. Data profiling is an interactive process between the user and the software. The results generated by Informatica Data Explorer are reviewed by the user, enabling them to arrive at a model that is both consistent with the source data and is meaningful in a business context.

Column Profiling: Profiles each column to discover critical information about the attributes, including frequency and distribution of data values, patterns, nulls, and min and max values.

Single-table Structural Profiling: Also referred to as dependency profiling, discovers and validates critical relationships in the data structure including key analysis and key validation.

Cross-table Structural Profiling: Allows for the automated comparison of values across tables to identify duplicate data or redundancies.

Data Viewing: Allows the user to view the rows of data being investigated.

Data Rules: Allows the user to test rules against the data.

Orphan Analysis: Test for parent/child relationships across tables.

Data Mapping

Informatica Data Explorer provides robust data mapping capabilities, which generate accurate source-to-target mapping between different data structures and defines the necessary transformation specifications. With these capabilities, Informatica Data Explorer enables data analysts and architects to compare actual data and metadata sources to target application requirements and identify the gaps, inconsistencies, redundancies, and inaccuracies so that they can be resolved before moving the data.

For new applications and data warehouses, Informatica Data Explorer assists in creating new database designs, including producing data definition language (DDL) that is fully compatible with the source systems. The product derives a correct normalized data model and allows the user to manipulate that model to suit specific business needs.

For existing applications that already have predefined schemas and requirements, Data Explorer's fixed target mapping (FTM) functionality captures the target data structures from packaged applications (or data warehouse designs) and manages the mapping process within the Data Explorer repository between profiled sources and the target. Data Explorer FTM can deliver matching specifications between multiple data structures. In addition Data Explorer FTM/XML mapping goes a step further, supporting applications that utilize XML to communicate with back-end databases and beyond a company's borders.

Knowledge is Power

The Informatica Data Explorer open repository stores information discovered during the profiling process and makes this information available to other components and products including Informatica Data Quality. Information in the Informatica Data Explorer repository is also used by Informatica

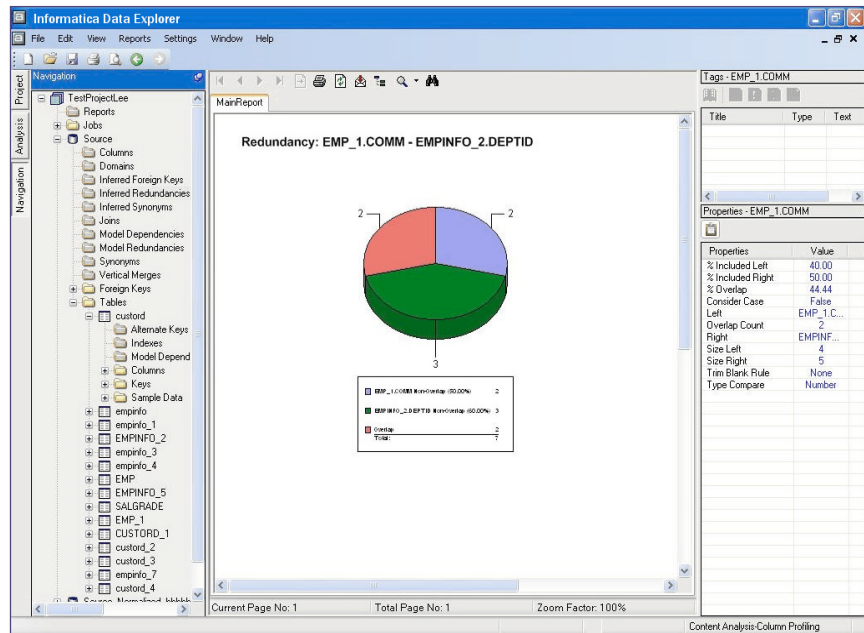


Figure 2: Informatica Data Explorer generates graphical reports quickly and easily.

PowerCenter to drive and speed-up the data integration process. At the same time data analysts and project teams can view the catalog of profiling information and produce graphical reports and dashboards to analyze the data. This rich repository of accurate information establishes an open, accessible and consistent method for sharing the source system knowledge discovered.

Integration with Informatica Data Quality and PowerCenter

Integration between Informatica Data Explorer and Informatica's other core data integration and data quality products is central to how the platforms work. Informatica Data Quality is a leading information quality solution that enables organizations to enhance, standardize and cleanse data based on easily configured business rules. It can use the information provided by Informatica Data Explorer to drive ongoing data quality strategies and deploy processes interactively, as batch processes or in real-time. The two

products are linked via an open repository, offering organizations a single platform for tackling data quality at multiple points across the organization while maintaining centralized control and management of data quality standards. The combination of Informatica Data Quality and Informatica Data Explorer provides businesses and government organizations with the industry's most complete platform for successful planning and execution of data quality initiatives.

Informatica Data Explorer is integrated with Informatica's leading data integration product PowerCenter via fixed target mapping (FTM). From within PowerCenter, a developer can use Informatica Data Explorer's profiling results to create data integration mappings. All automated profiling results are available to PowerCenter users including all relevant information discovered about data and data sources, which can dramatically speed-up and streamline the data integration process.

Summary

Informatica Data Explorer supports data investigation, discovery and mapping that can be used by the data analyst for data migration, data integration, and data quality projects. Its value lies in its ability to quickly analyze multiple source systems and assist the user in scoping the size of a migration, data integration or data quality project.

By quickly providing a complete, 100 per cent accurate picture of your data across all enterprise systems, Informatica Data Explorer helps ensure that there are no surprises in your data, enabling your organization's valuable IT initiatives to proceed on time, on budget and with reduced risk.

- Informatica Data Explorer automates the analysis of existing source data to discover its true content, structure and quality.
 - Informatica Data Explorer profiles all data types in all data sources, including flat files, relational databases, enterprise applications, mainframe VSAM, IMS.
 - Informatica Data Explorer provides a wealth of profiling information and metadata and makes it available to project teams, other Informatica products and third party products.
- Informatica Data Explorer features an easy-to-use user interface that supports the intuitive investigation and documenting processes that data analysts use.
 - Informatica Data Explorer generates accurate source-to-target mapping and produces transformation specifications.

By providing a complete set of data profiling and mapping capabilities in a single easy to use product, Informatica minimizes risk in data-centric projects by ensuring data is fit for its intended purpose and by eliminating unfounded assumptions



Data Quality Division, Wilson House, Fenian Street, Dublin 2, Ireland
phone: +353.1.4004900 fax: +353.1.4004999 www.informatica.com

Worldwide Headquarters, 100 Cardinal Way, Redwood City, CA 94063, USA
phone: 650.385.5000 fax: 650.385.5500 toll-free in the US: 1.800.653.3871 www.informatica.com

Informatica Offices Around The Globe: Australia • Belgium • Canada • China • France • Germany • Japan • Korea • the Netherlands • Singapore • Switzerland • United Kingdom • USA