

# Informatica: Making Data Work

## Enterprise-wide data quality with Informatica Data Explorer and Informatica Data Quality

### BENEFITS:

- Reduce cost, risk, and time required to implement new systems
- Increase accuracy of enterprise data assets, instilling confidence in data used by business
- Empower business users to manage and deploy data quality programs enterprise-wide
- Reduce opportunity costs of lost sales, increase customer satisfaction, and reduce shipping or mailing expenses due to poor quality data

Data quality is an increasingly serious issue for organizations large and small. It is central to all data integration initiatives. Before data can be used effectively in a data warehouse, or in customer relationship management, enterprise resource planning or business analytics applications, it needs to be analyzed and cleansed. To ensure high quality data is sustained organizations need to apply ongoing data cleansing processes and procedures, and to monitor and track data quality levels over time. Otherwise poor data quality will lead to increased costs, breakdowns in the supply chain and inferior customer relationship management. Defective data also hampers business decision making and efforts to meet regulatory compliance responsibilities.

As a business information professional you're responsible for making sure that the data upon which your company bases its strategic business decisions and processes is complete; conforms to standard formats, is consistent, accurate, and free from duplicates. You need technology designed specifically for you so you can design, manage, deploy, and control data quality solutions enterprise-wide. You need **Informatica Data Explorer** and **Informatica Data Quality**. These unique, easy-to-use data quality software products are specifically designed to bridge the gap between and better align IT and the business, to enable them to control data quality processes enterprise-wide in order to reduce costs, increase revenue, and improve business decision-making.

Informatica has the most comprehensive data quality solution available on the market. Our products, Informatica Data Explorer and Informatica Data Quality, offer all the functionality needed to initiate and manage data quality programs, from upfront profiling and investigation of multiple data sources through to cleansing, de-duplication, monitoring, validation and consolidation.

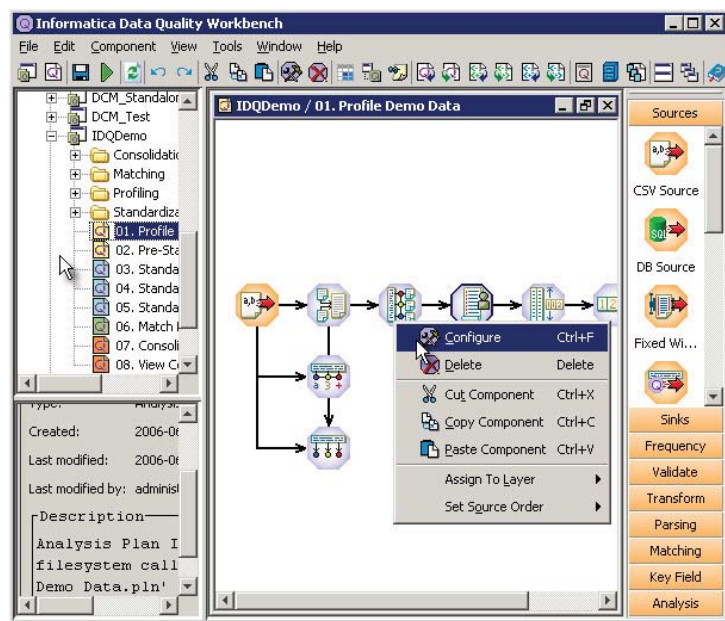


Figure 1: Informatica Data Explorer and Informatica Data Quality Workbench

## ANALYZE ALIGN CLEANSE SUSTAIN

The combination of Informatica Data Quality and Informatica Data Explorer provides organizations with the industry's most complete platform for successful planning and execution of one-off or ongoing data quality initiatives. Informatica Data Explorer enables the in-depth upfront investigation and profiling of multiple data sources, and Informatica Data Quality is the complete platform enabling business analysts to design, manage and deploy data quality processes across the enterprise. Together these best-in-class products provide a bridge between the IT and business professionals involved.

### Analyze

Data quality begins with understanding all the data in your source systems. Informatica Data Explorer is breakthrough software that puts powerful data profiling and mapping capabilities in the hands of the business and IT professionals who need them.

Informatica Data Explorer profiles data in three dimensions:

- Column profiling – analyzes all the values within each column or attribute. Discovers the TRUE metadata and points to data quality problems.
- Single-table structural profiling – examines each attribute in relation to every other attribute/column within a table, looking for dependency relationships. Discovers functional dependencies, primary keys and data structure quality problems.
- Cross-table structural profiling – compares data between tables, determining which attributes contain overlapping or identical sets of values. Discovers duplicate data across systems; foreign keys, synonyms, and homonyms; and values corrupting data integrity.
- All the information discovered about data sources is stored in a repository making it available for other products, including Informatica Data Quality for data cleansing and monitoring and Informatica PowerCenter for data integration, as well as third-party data management products.

### Benefits

- Automates data investigation, making the data profiling process faster and easier
- Allows for resolution of data quality issues before the data is cleansed or used
- Gives full responsibility for data investigation to those in the organization who know and understand the data best

### Align

Using the information gathered from the data profiling process, Informatica Data Explorer builds a normalized data model, in which unwanted redundancies are eliminated. The model can then be used as a staging area for movement to a fixed target, or as an operational data store.

Informatica Data Explorer is fully integrated with the Informatica platform through the Fixed Target Mapping capability (FTM). This FTM functionality captures the target data structures from packaged applications (or datawarehouse designs) and manages the mapping process between profiled sources and the target.

In addition, Informatica Data Explorer maps profiled data sources to targets such as eCommerce applications, data warehouses and packaged CRM and ERP systems.

The business and IT community can use the results of data profiling to communicate effectively, understand their business processes and discover where they are flawed. By using the information in the Informatica Data Explorer repository, they will ask the right questions, and will build better, more accurate plans and specifications.

### Benefits

- Increase effectiveness of new system implementations when metadata is accurate
- Increase productivity by enabling users to easily communicate data quality issues through tags
- Reduce risk of duplicating data issues in broader data integration initiatives through integration with PowerCenter

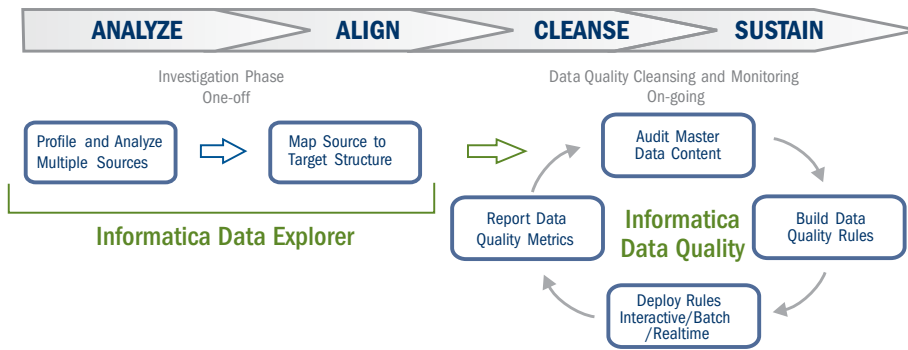


Figure 2: The PowerCenter Enterprise Grid Option expands the platform's capabilities to include Session on Grid and high availability, which enhance the platform's scalability and reliability when deployed on a grid computing environment, minimizing the risk of outages and reducing costs associated with data downtime.

## Cleanse

Once the data sources have been profiled, the next phase of the process uses that analysis to design and automate data cleansing strategies. This is where Informatica Data Quality steps in.

Informatica Data Quality is unique software that is specifically designed to put the control of data quality processes in the hands of the business. With unparalleled ease of use, the software delivers powerful data cleansing, matching, and reporting and monitoring capabilities in a single solution. Data analysts and data stewards use Informatica Data Quality to easily design, manage, deploy, and control individual and enterprise-wide data quality initiatives. By providing a complete process for measuring, monitoring, tracking, and improving data quality at multiple points across the organization over time, Informatica Data Quality empowers business information owners to implement and manage effective and lasting data quality processes enterprise-wide.

Informatica Data Quality is centered on the Informatica Data Quality Workbench, used for building data quality rules and managing reference dictionaries and scorecards. In combination with Informatica PowerCenter, Informatica Data Quality's platform also includes the functionality to deploy solutions in batch and realtime for ongoing cleansing, enhancement and data quality monitoring. Informatica Data Quality is fully integrated with Informatica PowerCenter via PowerCenter Mapping Designer and Workflow.

## Benefits

- Empowers business professionals to manage and deploy ongoing or one-off data quality programs for all data.
- Enables organizations to re-use data quality business rules, storing and deploying the same rules in all environments – interactively, as scheduled batch jobs or in real-time, saving time and reducing costs.

## Sustain

To date, many organizations that have attempted to tackle data quality have tended to implement tactical solutions to improve quality within a single application or within a single business process. While this approach may mitigate the problem for part of the organization in the short term, such limited initiatives generally fail to achieve long-term data quality improvement. To solve the data quality issue for good requires an enterprise-wide approach that includes addressing organizational, cultural, process and technology infrastructure.

This can be achieved on a phased basis, but for data quality improvement to become a reality, one from which your organization reaps dramatic benefits, the long term view must be taken. Maintenance, measurement and monitoring are central to all quality management systems, and data quality is no exception. Enterprises that systematically measure data quality provide a foundation for effective data quality improvement as well as for understanding the appropriate uses of data.

Informatica Data Quality supports ongoing data quality improvement through its iterative data quality management process and powerful reporting functionality. Designed for use by teams of data and business analysts, Informatica Data Quality enables users to quickly and easily develop customized rule sets for automated data quality reporting, validation and correction. These rules can then be deployed in interactive, batch or real-time environments to ensure data quality is maintained and improved over time.

The product provides management-level scorecards that monitor the six key attributes of data quality (completeness, conformity, consistency, accuracy, duplication and integrity) across all data. The drill-down reports enable users to zoom in from high-level views to inspect low-quality data records one-by-one and to identify issues in an iterative discovery and cleanse process.

Informatica Data Quality supports and drives methodologies for ongoing data quality management and improvement.

## Benefits

- Increase accuracy of enterprise data assets, instilling confidence in data used by business
- Empower business users to manage and deploy data quality programs enterprise-wide by tracking data quality improvements against targets.

## Conclusion

Is your organization involved in any type of data migration, data integration or data consolidation initiative that relies on an accurate understanding of the data sources? Are you implementing ERP, CRM, eCommerce, business analytics or Data Warehousing initiatives or merging data as a result of a merger or acquisition? Have you witnessed poor quality, corrupt, invalid, inconsistent, redundant and inaccurate data within your corporate systems? If yes, then you need Informatica. Why? Because Informatica can help you analyze, align, cleanse and sustain your corporate data to minimize risk in data integration projects and to drive ongoing data quality and regulatory compliance.

**INFORMATICA**  
The Data Integration Company™

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](http://www.informatica.com)

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

© 2008 Informatica Corporation. All rights reserved. Printed in the U.S.A. Informatica, the Informatica logo, Data Quality, and Data Explorer are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be trade names or trademarks of their respective owners.

## ABOUT INFORMATICA

Informatica Corporation is a leading provider of enterprise data integration software. Using Informatica products, companies can access, discover, cleanse, integrate, and deliver enterprise data across systems, processes, and people to reduce complexity, ensure consistency, and empower the business.