

# DataFlux Data Management Platform

The DataFlux Data Management Platform provides the key for true business and IT collaboration – a single solution for data management initiatives.

## A Single Point of Control for Enterprise Data Management

Organizations today are faced with a daunting challenge: how to control the information that serves as the very foundation of their business success. However, with the recent exponential growth in data and the proliferation of siloed, disparate data, organizations are realizing that corporate information often cannot support the operational and analytical systems that drive the business.

The DataFlux Data Management Platform provides a unified solution for both business and IT users to plan, implement and monitor business-critical information. The Data Management Platform helps organizations establish and maintain a single, consistent set of data policies and processes for all corporate data. This enables organizations to more effectively establish and enforce uniform standards for enterprise data – a critical component for data governance, master data management (MDM) and other enterprise data management initiatives.

By delivering a single, accurate and consistent view of DataFlux tasks through a user-friendly process and technology framework, the Data Management Platform lets you:

- Enable business and IT users with both web-based and traditional client interfaces for core data management tasks
- Enhance cross-functional data governance through a single platform to create, optimize and manage business rules
- Analyze, monitor and actively manage the quality of any enterprise data, in any application, across the organization

## Data Management at Your Fingertips

The Data Management Platform offers a unique set of technologies built on an industry-leading technology platform that encompasses every facet of the data management process. The Data Management Platform provides business and IT users with powerful data improvement capabilities and control over data management and data governance initiatives.

The Data Management Platform provides the ability to design and manage processes to:

- Merge customer, product or other enterprise data
- Unify disparate data through a variety of data integration methods (batch, real-time, virtual)
- Verify and validate customer and product information
- Integrate disparate data sets, with an eye on quality
- Transform and standardize various data entities
- Monitor data for compliance and data quality control
- Manage metadata hierarchies

By enabling consistent, accurate and timely data throughout the enterprise, the Data Management Platform gives organizations the ability to implement the people, process and technology changes necessary to establish an effective data governance strategy. The Data Management Platform provides powerful functionality to meet today's demanding data management challenges.

# Technology for Data Stewardship and Data Governance

DataFlux web-based technology powers data governance initiatives and gives data stewards control of critical data management tasks. Effective data governance strategies allow companies to manage corporate information and make the right business decisions to stay competitive – and thrive in an uncertain business environment.

## DataFlux Business Data Network

Create a “common language” for business data

DataFlux Business Data Network gives business and IT staff a web-based interface to collaborate on the creation and management of business terms. Users can create descriptions and associated requirements of a term, helping give the technical team guidance on how to create business rules in the DataFlux Data Management Studio application. Business Data Network provides access, sharing and exchange of terms and attributes by providing views of:

- Descriptions
- Source systems
- Owners
- Related terms and processes (data management services, data workflows and applications)

By linking terms to business rules and data monitoring processes created in the Data Management Platform, Business Data Network provides an important starting point for data governance efforts. Data stewards, IT staff and enterprise architects can leverage the terms to align proper data across projects or business units. So, whether your staff is located in the next office or in another country, you have a consistent set of terminology driving data management strategies.

The screenshot displays the DataFlux Business Data Network interface. At the top, it shows 'Business Data Network' and 'DataFlux Web Studio'. Below the header is a navigation bar with icons for 'New Term...', 'Open...', 'Tags...', and 'Monitor'. A table lists several terms with columns for Term Name, Description, Importance (represented by stars), Status, and Last Modified. The 'Region' term is selected and highlighted in blue. Below the table, the 'Region' term's details are shown, including a description, requirements, attributes, and related terms.

Term Name	Description	Importance	Status	Last Modified
Legal	Legal protects the legal rights an...	★★★★★	Under R...	10/4/2011 8:02:22 AM
Marketing	Marketing at D&G stores is desin...	★★★★★	Production	9/23/2011 11:22:37 AM
Merchandising	Merchandising plan what is goin...	★★★★★	Production	9/23/2011 10:14:54 AM
Operations	Operations is responsible for the...	★★★★★	Production	9/23/2011 10:14:40 AM
Region	Area of the country which contain...	★★★★★	Production	10/4/2011 7:46:10 AM
Technology	The IT Department provides cost...	★★★★★	Production	9/23/2011 9:39:39 AM

**Region**

**Description**  
Area of the country which contains a group of stores.

**Requirements**  
Valid regions are represented in the Reference Data Manager. System checks for valid region codes must be made against the Region Codes lookup table. Regions must be in the United States. Updates to regions, region codes, and arrangement of stores to region must be approved by the Data Governance council before changes are pushed to production systems

**Attributes**  
Primary Role: Grouping

**Details**  
Status: Production  
Importance: Critical  
Locked by: Todd Haley

**Links**  
Region Map Relationships

**Tags**  
Data Warehouse

**Related Terms**  
Customer  
Seasonal

*Business Data Network allows users to collaborate on the creation of business terms, establish links to the term in a hierarchy of reference data, and examine related data elements.*

## DataFlux Reference Data Manager

*Build and manage relationships and hierarchies across data elements*

DataFlux Reference Data Manager offers a web-based interface that can be used across departments and lines of business to store and manage different types of hierarchies. This technology uses terms created in Business Data Network and other data elements to link relationships across data elements, including parent-child, one-to-many or coded relationships.

Hierarchy information can create a more complete view of the relationships that these business terms have to one another. Reference Data Manager allows data governance and data management teams to:

- Maintain reference data entities consistently across the organization, with definitions and relationships clearly defined for each entity and the entity values
- Identify, profile and validate reference data from any source system
- Create reference data required for effective master data management efforts

## DataFlux Web-based Dashboards and Monitoring

*Track the performance and health of your data management project*

One of the key facets of an effective data governance program is the continuous feedback of information to a data governance council and board – as well as the management structure behind the effort. DataFlux provides web-based technology that delivers a lightweight interface for management, data stewards and other interested parties to view critical metrics to gain insight into the health of corporate data.

DataFlux technology provides the foundation for data governance teams to:

- View data alerts to show when data is out of compliance with established data policies
- Create customized dashboards for individual users based on data types, data sources or other aspects of the data management program
- Deliver real-time information on data integrity to all members of a data stewardship team

# Technology for Business and IT Collaboration

DataFlux technology provides a powerful client interface – DataFlux Data Management Studio – that allows business and IT users to collaborate on the core components of data management. From initial data discovery through business rules development and monitoring, DataFlux provides a single interface to create, refine and implement a cohesive plan for data management efforts across the enterprise.

## Metadata Analysis

*Extract, organize and analyze metadata anywhere*

Metadata analysis uncovers existing trends and characteristics of your data, examining corporate data throughout your enterprise and providing a clear picture of the types and sources of that data. This understanding is an essential first step to any enterprise data management initiative. DataFlux provides the ability to discover and manage metadata from virtually any data source – anywhere in the organization – from a single interface.

## Data Profiling

*Discover data characteristics to guide data quality and data integration efforts*

Data profiling provides an in-depth assessment of your organization's data, examining the structure, completeness and suitability of your data assets. DataFlux provides industry-leading data profiling capabilities, providing insight into the health of your data. Data profiling is critical when you need to identify the root cause of poor-quality and disparate data sources. Through a data discovery program, Data Management Studio helps you gain the knowledge to design effective data quality, data integration and master data management business rules to support your organization's data-driven initiatives.

*Data profiling provides immediate feedback on the accuracy and integrity of data sources.*

The screenshot displays the DataFlux Data Management Studio 2.2 interface. The main window shows a 'Pattern Frequency Distribution' table for the 'Zip' field. A secondary window is open for the 'Post\_Code' field, also showing a 'Pattern Frequency Distribution' table. The interface includes a menu bar, a toolbar, and a left-hand navigation pane with sections for 'Collections', 'Tables', and 'Redundant Data Analyses'.

Pattern	Alternate	Count	Percentage
999999999	% (9)	41479	84.03
99999	% (5)	7679	15.56
99999-9999	% (5)-% (4)	204	0.41
9999	% (4)	2	0.00
99999-999	% (5)-% (3)	1	0.00

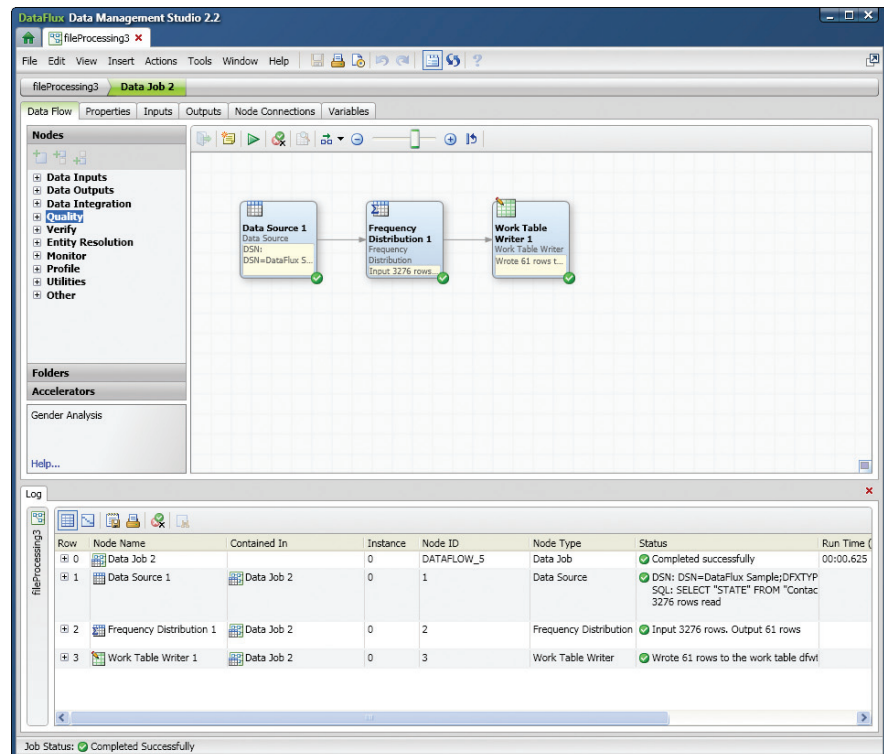
Pattern	Alternate	Count	Percentage
999999999	% (9)	37599	76.09
99999	% (5)	6698	13.56
99999-9999	% (5)-% (4)	4962	10.04
9999	% (4)	104	0.21
99999	% (5)	31	0.06
999	% (3)	4	0.01
999	% (3)	3	0.01
999	% (3)	2	0.00
99 9999	% (2) % (4)	1	0.00
999999	% (3) % (3)	1	0.00
AA99 9AA	A(2)P(2) 9A(2)	1	0.00
99999-999	% (5)-% (3)	1	0.00
aa9 99a	a(2)9 9(2)a	1	0.00
AA9 9A9	AA9 9A9	1	0.00
9999999	% (7)	1	0.00
AA99999 999 AA999	Aa(6) a(3) Aa(4)	1	0.00
999999	999999	1	0.00

## Data Quality

*Standardize, rationalize and transform corporate information*

Better data leads to better decisions which, ultimately, lead to better business. Data Management Studio gives both business and IT users the full capabilities of the DataFlux industry-leading data quality technology. Through patented matching technology, transformation routines and identification logic, DataFlux helps you easily correct data problems for virtually any and all types of enterprise data.

*A collaborative design environment allows business and IT users to develop and refine data jobs and services.*



The screenshot displays the DataFlux Data Management Studio 2.2 interface. The main workspace shows a workflow with three nodes: 'Data Source 1' (Data Source), 'Frequency Distribution 1' (Frequency Distribution), and 'Work Table Writer 1' (Work Table Writer). The log window at the bottom shows the following details:

Row	Node Name	Contained In	Instance	Node ID	Node Type	Status	Run Time
0	Data Job 2		0	DATAFLOW_S	Data Job	Completed successfully	00:00.625
1	Data Source 1	Data Job 2	0	1	Data Source	DSN: DSN=DataFlux S... SQL: SELECT "STATE" FROM "Contac 3276 rows read	
2	Frequency Distribution 1	Data Job 2	0	2	Frequency Distribution	Input 3276 rows. Output 61 rows	
3	Work Table Writer 1	Data Job 2	0	3	Work Table Writer	Wrote 61 rows to the work table dfwt	

Job Status: Completed Successfully

## Data Integration

*Intelligently match, merge and consolidate data*

An effective data integration strategy can lower costs and improve productivity by enabling consistent, accurate and reliable data across your enterprise. Data Management Studio provides users with a single interface for data quality and data integration activities, including the design, navigation and management of data integration jobs and workflows. Data integration involves combining processes and technology to enable your enterprise to make the most effective use of disparate, inconsistent data.

## Data Monitoring

*Enforce business rules and build the foundation for data governance*

By creating business rules once and reusing them across applications, you can apply a uniform set of business standards in real time across any system. Data Management Studio provides the design, development and monitoring environment for proactive data governance, so you can design and enforce rules to determine if data is maintained within proper control limits and meets predefined business rules.

## Data Enrichment

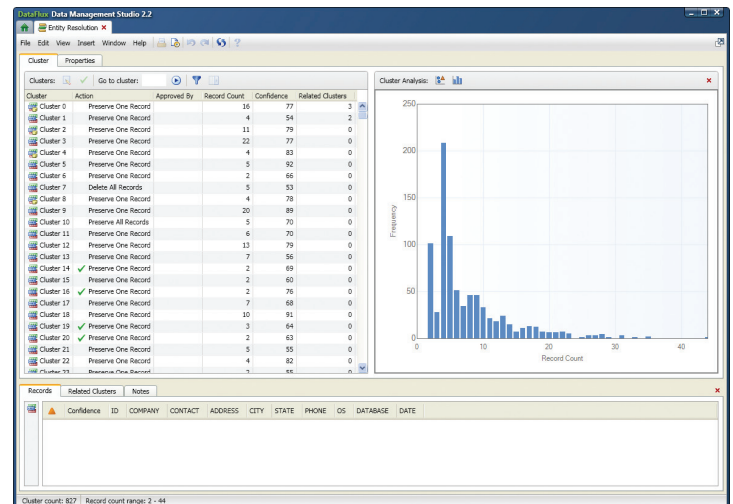
*Transform incomplete data into useable, standardized information*

DataFlux technology enables data verification and standardization – key components in data quality projects that contain customer, supplier, company or employee data elements. DataFlux gives you the ability to create complete and accurate address information for more than 240 countries and territories around the world. In addition, DataFlux can enrich address data with geographic, demographic or other details, as well as standardize and augment data on products, materials and services.

## Entity Resolution

*Match data and identify potential data relationships across sources*

The ability to link and consolidate entity information with a high level of confidence is critical to data management initiatives. Information about the same customer, product or employee may exist in multiple databases, in many unique forms. The challenge is to find and resolve similar records in different data sources. DataFlux offers industry-leading matching technology that enables accurate entity resolution across multiple data sources.



*Entity resolution allows users to review and manage identified data relationships.*

## Data Management Foundations

*Rapidly build and deploy a master data hub*

MDM efforts are often part of a larger data management or data governance initiative within the organization. With DataFlux technology, you can configure a pre-built master data model within Data Management Studio, providing the foundation for accelerated MDM deployments. By creating a master hub within DataFlux technology, you can tie a master data effort more closely to existing data management initiatives – and seamlessly move to DataFlux qMDM Solution for advanced MDM capabilities.

# Technology for Enterprise Deployment

---

*DataFlux servers allow you to extend the rules and services of Data Management Studio across multiple data sources throughout the enterprise. These robust servers allow organizations to deploy a single set of data management rules and workflows in batch, real-time or virtual deployments.*

## DataFlux Data Management Server

*Extend data management across the IT infrastructure*

Working in unison with Data Management Studio, DataFlux Data Management Server helps form the backbone of the Data Management Platform. The technology can implement the rules created in Data Management Studio in both batch and real-time environments.

Through an innovative, standards-based service-oriented architecture (SOA), Data Management Server frees you from writing and testing hundreds of lines of extra code to enable real-time data quality, data integration and data governance routines. Data Management Server is the workhorse that enables pervasive data quality, data integration and MDM throughout your organization. Data Management Server can:

- Enable the complete reuse of data quality and integration business logic
- Schedule and execute batch and real-time jobs to maximize server performance
- Integrate DataFlux data quality and data integration with existing applications to move data management routines further upstream
- Expose jobs as web services through BPM/BPEL integration

## DataFlux Federation Server

*Use virtual data views to rapidly access and view enterprise data*

DataFlux Federation Server provides integrated views of data located in multiple repositories without the need to physically consolidate the data. This virtual integration enables powerful querying capabilities, as well as improved source data management. Data federation efficiently joins data from many heterogeneous sources, without moving or copying the data. The result is improved performance and speed, while reducing costly and labor-intensive data movement and consolidation.

DataFlux Federation Server enables organizations to view data from multiple sources through an integrated, virtual data view. While the data remains stored in its source application, multiple users can see integrated data – without physically moving or changing data. Federated data capabilities available in DataFlux Federation Server include:

- A central location for setup and maintenance of database connections
- Access to common database systems, including Oracle, DB2, Teradata, SQL Server and MySQL
- TCP/IP server access
- ODBC and native drivers to connect to data sources
- Permissions that allow only specific users to access and control data



Your Data.  
Your Business.  
One Solution.

[www.dataflux.com](http://www.dataflux.com)

**Corporate Headquarters**

940 NW Cary Parkway  
Suite 201  
Cary, NC 27513  
877 846 3589 (USA & Canada)

**DataFlux West**

15300 Ventura Blvd.  
Suite 523  
Sherman Oaks, CA 91403  
818 906 7638 (USA & Canada)

[info.us@dataflux.com](mailto:info.us@dataflux.com)