

# Informatica Data Quality





## Agenda

- Data Quality Overview
- Release 9.1 Architecture Overview
- Informatica Analyst
- Address Validation Overview
- Data Quality Matching
- PowerCenter Integration



# Informatica Data Quality

## Data Quality Overview

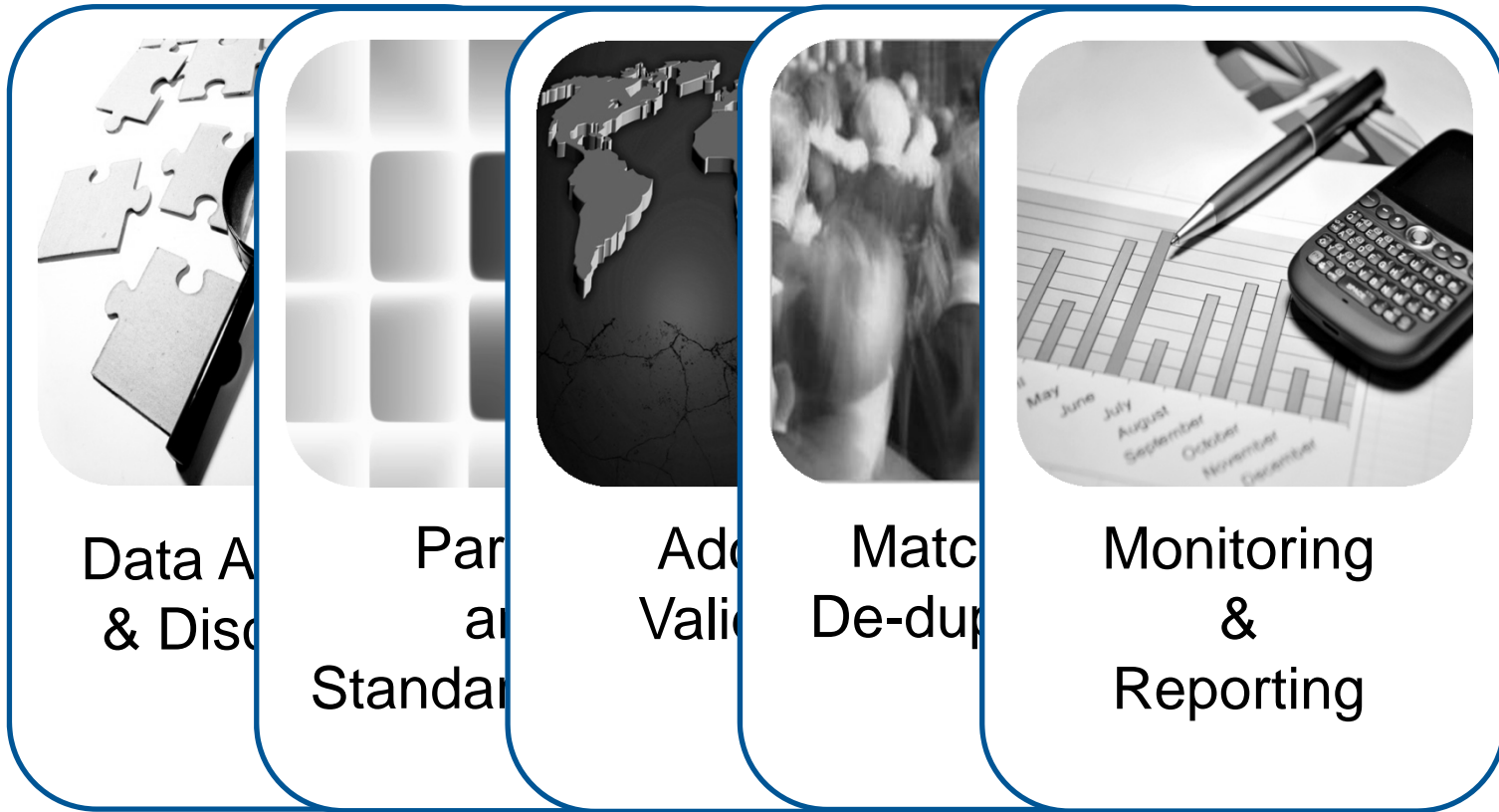


# Informatica Data Quality

Informatica Data Quality provides...

*Powerful analysis, cleansing, matching, exception handling, reporting and monitoring capabilities that enable IT and the business to manage enterprise-wide data quality initiatives.*

# Frequent Requirements



And do this for all data types...

# Data Analysis and Discovery



# Web Based Analyst Tools to Profile Your Data



Data  
Steward

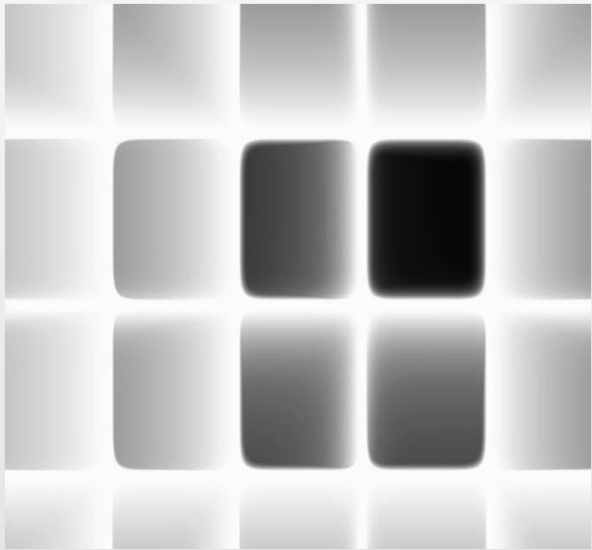
The screenshot shows the Informatica Analyst web interface. The main window displays a 'Column Profiling' report for a project named 'Profile\_Customer\_Cor...'. The report is organized into three tabs: 'Values', 'Patterns', and 'Statistics'. The 'Patterns' tab is currently selected, showing a table of patterns with columns for 'Pattern', 'Frequency', and 'Percent'. The patterns listed are 'Others' (Frequency: 21, Percent: 11.86%), '9999' (Frequency: 20, Percent: 11.30%), and '99999' (Frequency: 136, Percent: 76.84%).

Pattern	Frequency	Percent	Chart
Others	21	11.86	
9999	20	11.30	
99999	136	76.84	

Below the patterns table, a data preview table is visible, showing columns for CUSTOMERTIER, NAME, ADDRESS1, ADDRESS2, CITY, STATE, and ZIP. The data rows include customer information such as Silver, Diamond, Bronze, and Ruby tiers, along with their respective names and addresses.

Increase productivity and efficiency by enabling the business to proactively take responsibility for data quality and reduce their reliance on IT.

# Parsing & Standardization



The key objectives are:

- to transform and parse data to multiple fields
- to correct completeness, conformity, and consistency problems
- to standardize field formats



# Parsing & Standardization: All Data

Product ID		Brand		Description
90017		iPod		4GB, Red iPod Nano //Special Edt.
Product_ID	Brand	Size	Color	Description
90017	IPOD	4GB	Red	4 Gigabyte Nano Special Edition (Red)

OrgName		CountryCode	Currency
AAVID THERMAL TECH (ATT)			US Dollar
LegalName	ShortName	CountryCode	Currency
Aavid Thermal Technologies Inc.	ATT	US	USD

ContactName				Phone
Judy Dent // Bob's Assistant				415.555.1212
FirstName	MiddleName	LastName	Title	Phone
Judy		Dent	Bob's Assistant	+1 (415)555-1212

# Address Validation

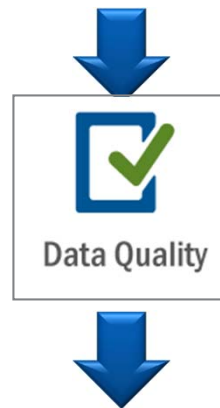


- Validate or correct addresses for over 240 countries
- Have reference data from international postal agencies
- Validate worldwide data in one environment
- Be continuously maintained with worldwide post offices and databases



# Address Validation

Address1	Address2	Address3	Address4	Address5
7887 KATY FRWY	SUITE 333	HUSTEN	TX	99999



Street	City	County	StateCode	StateName	ZIP	ZIP4	Latitude	Longitude
7887 Katy Freeway Suite 333	Houston	Harris	TX	Texas	77024	2005	29.283427	-95.46802

Valid addresses keep costs down and helps ensure compliance

# Match and De-Duplicate All Data Types



Using multiple attributes using multiple rule sets

Use confidence levels to automate

Consider over 60 cultural variations for name matching

Match data in spite of poor quality

# Match and De-Duplicate All Data Types

SKU	Description	Size	Price
AP-2199	Sailors' Desk Lamp	12 in	27.99
AP2199	Nautical Lamp	12 inch	27.99
PA-2119	Sailing Lamp	1 Foot	34.99

Intrinsically wrong (and potentially uncorrectable) data can still be valuable for Matching purposes

- Alternate or Nicknames
- Misspellings
- Invalid Data

Name	DOB	Address	City	State	Zip
W. S. Harrison II PhD	1/33/1967	Medical Center,117/2A #17497 Jackson	E. Hartford	NY	16987
William Stuart Harison	1/3/1967	117- 2a Jacksen Rd.	Easthartford	CT	06987
William Stewart Harison	9/9/99	117 Jackson Road. Suite 2A	Hartford East	CT	06987
Doctor Bill Harisen jr	1/13/1967	117 Jacson Room 2a	HartfordCT		6984
Harrisen William Doctor		2a Jackson Rd #174978	Hartford	CT	06987-4573

Highly accurate matching ensures the minimum number of duplicates

# Monitoring Quality



## Stakeholders need to be aware

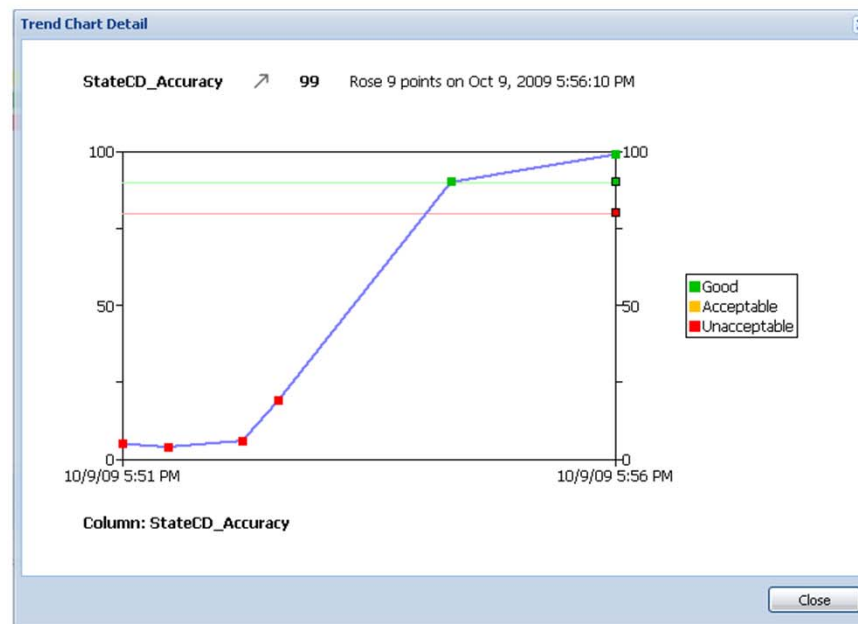
- Current quality metrics
- Alerts if quality thresholds are not being met

**Delivery of reports and alerts must be web based**

# Monitoring the quality

Easy-to-share browser-based scorecards for line of business managers

Scorecard_Before: Scorecard		
NAME ▲	SCORE	% VALID CURRENT
CustomerNum_Validity	87	<div style="width: 87%;"></div>
StateCd_Validity	90	<div style="width: 90%;"></div>
ZipCD_Accuracy	80	<div style="width: 80%;"></div>



Line of business manager

- Browser-based scorecards enabling you to:
- View and share data quality scorecards
- Drilldown to the actual records
- Take action to reduce the business impact

Zero learning curve for business users to review and track data quality metrics, enabling data quality “for the masses”.

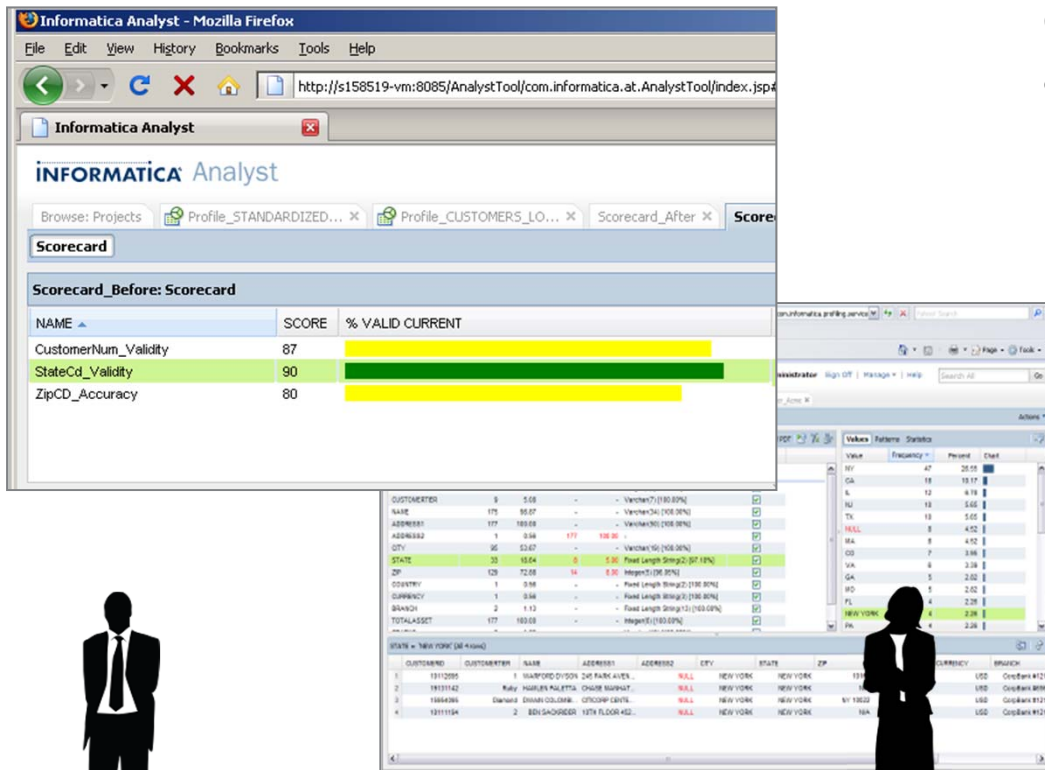


# Business Empowerment

## Simple-to-use browser-based tools

Designed for the tasks and skills of business data stewards and analysts

- Purpose-built, web-based UI for fast ramp-up
- Scorecarding & trending
- View business, not technical, representations
- Interact with data directly through profiling, rule validation, and scorecarding



  
**Business  
Manager**

  
**Analyst  
& Steward**



Work with relevant data to meet business needs while reducing reliance on IT



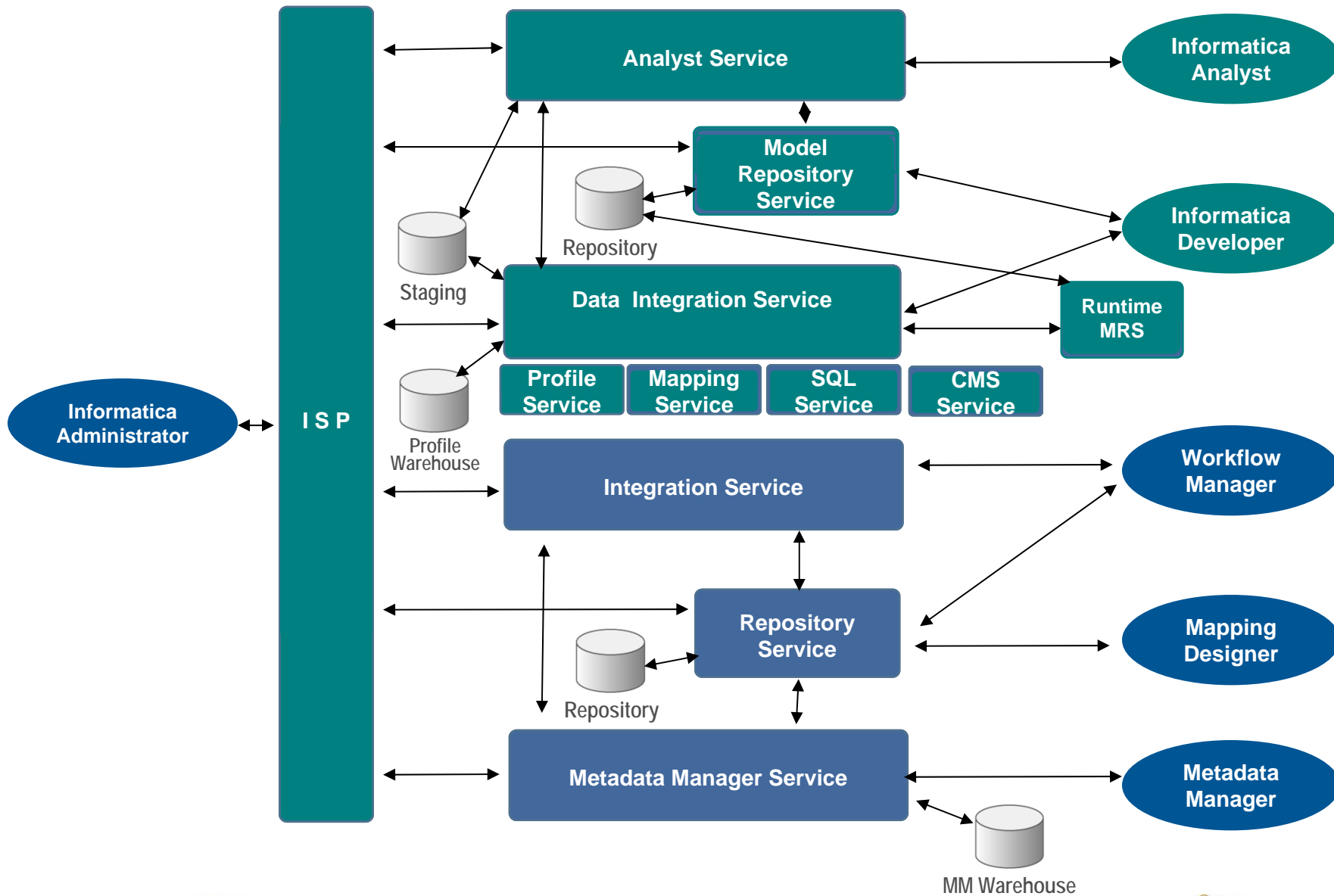


## Release 9.1 Architecture Overview

IDQ V9.1



# Release 9 DQ Architecture



# Model Repository Service (MRS)

- **Application service that manages the Model Repository**
  - Model Repository is a relational database that stores the metadata for projects created in the Analyst and Developer Tools
  - Stores run-time and configuration information for deployment units that are deployed to a Data Integration Service
- **Node configuration**
  - MRS can run on one node but is not a highly available service
  - Multiple MRS Services can run on the same node
  - If MRS fails, it automatically restarts on the same node

# Data Integration Service (DIS)

- ***Container*** for all data integration functionalities
- **Plug-ins provide different DI functionalities**
  - Profiling Service plug-in translates profile into mappings
  - SQL Service plug-in translates SQL into mappings
  - Mapping Service executes data quality plans
- **Provides common services to its plug-ins**
  - Request dispatch, thread pooling etc.
- **Mapping execution using embedded Data Transformation Machine (DTM - Interprets and executes mappings)**

# Analyst Service (AS)

- **Manages the Analyst tool**
- **Manages the connection to the directory that stores uploaded flat files that are used as flat file sources in the Analyst tool**
- **Manages the connection to a staging database that stores reference tables that are created or imported in the Analyst tool**
- **Connects to a DIS and MRS**

# Content Management Service (CMS)

- **Allows central configuration of content related items and provides content information to Informatica Developer.**
- **CMS contains configuration of Address Validation Parameters**
- **CMS is requirement for Identity Populations (used in Matching Transformation)**

# Informatica Services Platform (ISP)

- **Administrative Framework for Informatica Services**
  - Monitoring, start/stop and failover of services
  - Extensible via plug-ins for additional services
- **Supports single point of administration**
  - User, groups and role management
  - Service configuration and monitoring
- **Integration with LDAP, Active Directory**
- **Provides a set of core services used internally**
  - Authentication Service, Name Services etc.



# Informatica Analyst





# What is the Analyst 9.1?

## *Product Features/Highlights*

- **Data Sources**
  - Metadata import for Data Sources
  - Data access and preview
- **RTM (Reference Table Manager)**
  - Authoring and editing of reference data
  - Auditing of changes
- **Profiling**
  - Column Profiling
  - Rule Profiling
  - Rule creation/editing
- **Data Quality Scorecarding**
  - Create/Edit from just browser
  - Based on profile value frequencies
- **Collaboration via**
  - Metadata Bookmarks (URLs)
  - Profile Comments
  - Bi-directional Sharing of Rules with 9.1 Developer
  - Sharing of RTM dictionaries
- **DQA (Data Quality Assistant)**
  - Management of Bad Records and Duplicate Records
  - Auditing of changes
- **Mapping Specification**

Define business logic that populates a target table

Configure the sources, target, rules, filters, and joins to transform the data

# Profiling

## View

- Profile Statistics
- Value / Pattern Frequency Analysis
- Drilldown Analysis

INFORMATICA Analyst Administrator Log Out | Manage | Help Search by Name Go

Browse: Projects prf\_EMPLOYEES x

Column Profiling Data Preview Properties Actions

prf\_EMPLOYEES - Column Profiling Filter... Sampling...

Name	Uniqu...	Uniqu...	NULL	NULL %	Datatype	Inferr...	Documented C
Source Name							
EMPLOYEEID	18	100.00	-		Integer(2)	100.00	number(15)
LASTNAME	9	50.00	-		String(9)	100.00	varchar2(20)
FIRSTNAME	9	50.00	-		String(8)	100.00	varchar2(10)
TITLE	4	22.22	-		String(24)	100.00	varchar2(30)
TITLEOF COURTESY	4	22.22	-		String(4)	100.00	varchar2(25)
BIRTHDATE	9	50.00	-		Date Time	100.00	date(19)
HIREDATE	8	44.44	-		Date Time	100.00	date(19)
ADDRESS	9	50.00	-		String(30)	100.00	varchar2(60)
CITY	5	27.78	-		String(8)	100.00	varchar2(15)
REGION	2	11.11	8	44.44	Fixed Length Stri...	100.00	varchar2(15)
POSTALCODE	9	50.00	-		String(7)	100.00	varchar2(10)

Values Patterns Statistics

Value	Fre...	Per...	Chart
SALES REPRESENTA...	12	66.67	
VICE PRESIDENT, S...	2	11.11	
SALES MANAGER	2	11.11	
INSIDE SALES COO...	2	11.11	

Drill down: TITLE = 'SALES REPRESENTATIVE' (All 12 rows)

EMPLOYEEID	LASTNAME	FIRSTNAME	TITLE	TITLEOF COURTE	BIRTHDATE	HIREDATE	ADDRESS	CITY	REGION	POST
1	DAVOLIO	NANCY	SALES REPRESENTATIVE		08-Dec-2048 00:00:00	01-May-1992 00:00:00	507 - 20TH AVE. 1 APT. 2A	SEATTLE	WA	98122
3	LEVERLING	JANET	SALES REPRESENTATIVE		30-Aug-1963 00:00:00	01-Apr-1992 00:00:00	722 MOSS BAY BIKIRKLAND		WA	98033
4	PEACOCK	MARGARET	SALES REPRESENTATIVE		19-Sep-2037 00:00:00	03-May-1993 00:00:00	4110 OLD REDMCREDMOND		WA	98052
6	SUYAMA	MICHAEL	SALES REPRESENTATIVE		02-Jul-1963 00:00:00	17-Oct-1993 00:00:00	COVENTRY HOUSE MINER RD.	LONDON		EC2 7
7	KING	ROBERT	SALES REPRESENTATIVE		09-May-1969 00:00:00	03-Jul-1994 00:00:00	EDGEHAM HOLLOW	LONDON		EC2 7

# Profiling – Apply Rules

- Apply existing rule or create new one.
- View Profile results for Rule output

**INFORMATICA Analyst** Administrator Log Out | Manage | Help | Search by Name Go

Browse: Projects **prf\_EMPLOYEES**

**Column Profiling** Data Preview Properties Actions

**prf\_EMPLOYEES - Column Profiling** Filter... Sampling...

Name	Uniqu...	Uniqu...	NULL	NULL %	Datatype	Inferr...	Docu
CITY	5	27.78	-	-	String(8)	100.00	varch
REGION	2	11.11	8	44.44	Fixed Length Stri...	100.00	varch
POSTALCODE	9	50.00	-	-	String(7)	100.00	varch
COUNTRY	2	11.11	-	-	String(3)	100.00	varch
HOMEPHONE	9	50.00	-	-	String(14)	100.00	varch
EXTENSION	9	50.00	-	-	Integer(4)	100.00	varch
PHOTO	N/A	-	-	-	N/A	-	N/A
NOTES	9	50.00	-	-	-	-	varch
REPORTSTO	3	16.67	2	11.11	Integer(1)	100.00	numb
PHOTOPATH	5	27.78	-	-	String(34)	100.00	varch

**Rule: rule\_ValidateZip**

USZipValidity	2	11.11	-	-	String(7)	100.00	string
---------------	---	-------	---	---	-----------	--------	--------

**Values** Patterns Statistics

Value	Fre...	Per...	Chart
VALID	10	55.56	
INVALID	8	44.44	

Drill down: USZipValidity = 'INVALID'(All 10 rows)

EMPLOYEEID	LASTNAME	FIRSTNAME	TITLE	TITLEOFCOURTE	BIRTHDATE	HIREDATE	ADDRESS	CITY	REGION
2	FULLER	ANDREW	VICE PRESIDENTDR.		19-Feb-1952 00:00:00	14-Aug-1992 00:00:00	908 W. CAPITAL TACOMA	WA	
1	DAVOLIO	NANCY	SALES REPRESENTATIVE		08-Dec-2048 00:00:00	01-May-1992 00:00:00	507 - 20TH AVE. I APT. 2A	SEATTLE	WA
3	LEVERLING	JANET	SALES REPRESENTATIVE		30-Aug-1963 00:00:00	01-Apr-1992 00:00:00	722 MOSS BAY BIKIRKLAND	WA	
4	PEACOCK	MARGARET	SALES REPRESENTATIVE		19-Sep-2037 00:00:00	03-May-1993 00:00:00	4110 OLD REDMCREDMOND	WA	
8	CALLAHAN	LAURA	INSIDE SALES CLERK		09-Jan-1958 00:00:00	05-Mar-1994 00:00:00	4726 - 11TH AVE. SEATTLE	WA	

# Scorecards

- Add Profile Columns to a Scorecard
- Scorecard Name, Location Description.
- Set Thresholds

The screenshot shows the Informatica Analyst interface. The main window displays 'prf\_EMPLOYEES - Column Profiling' with a table of data. A dialog box titled 'Add to Scorecard: Step 3 of 3' is open, showing the configuration for a score named 'USZipValidity' based on the 'USZipValidity' column. The dialog includes a 'Score Settings' section with custom thresholds and a 'Score using: Values' section with a table of values and frequencies.

Name	Uniqu...	Uniqu...	NULL	NULL %	Datatype	Inferr...	Docu
CITY	5	27.78	-	-	String(8)	100.00	varch ^
REGION	2	11.11	8	44.44	Fixed Length Stri...	100.00	varch

Value	Frequency	Is Valid
VALID	10	<input checked="" type="checkbox"/>
INVALID	8	<input type="checkbox"/>

Score Name	Column Name
USZipValidity	USZipValidity

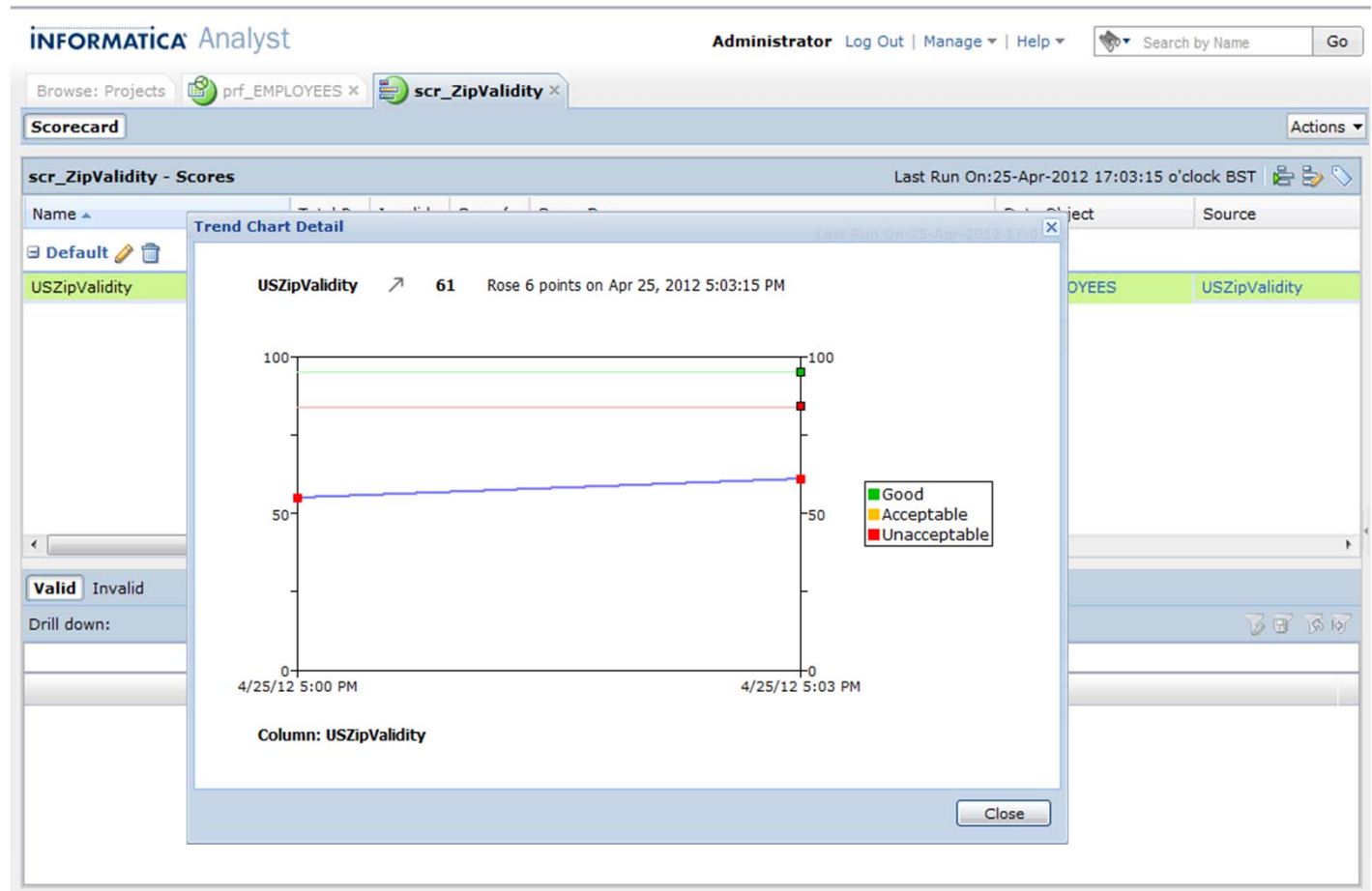
Score Settings: USZipValidity
Score Thresholds: <input checked="" type="radio"/> Set Custom Thresholds for this Score
95 to 100 % <span style="color: green;">■</span> Good
85 to 94 % <span style="color: yellow;">■</span> Acceptable
0 to 84 % <span style="color: red;">■</span> Unacceptable

ADDRESS	CITY	REGION
0:(908 W. CAPITAL TACOMA	WA	
0:(507 - 20TH AVE.   APT. 2A	SEATTLE	WA
0:(722 MOSS BAY BIKIRKLAND	WA	
0:(4110 OLD REDMCREDMOND	WA	
0:(4726 - 11TH AVE.SEATTLE	WA	

# Scorecards

- Run the Scorecard
- View Trend Chart



## Address Validation Overview



# V9 Data Quality Content Packaging

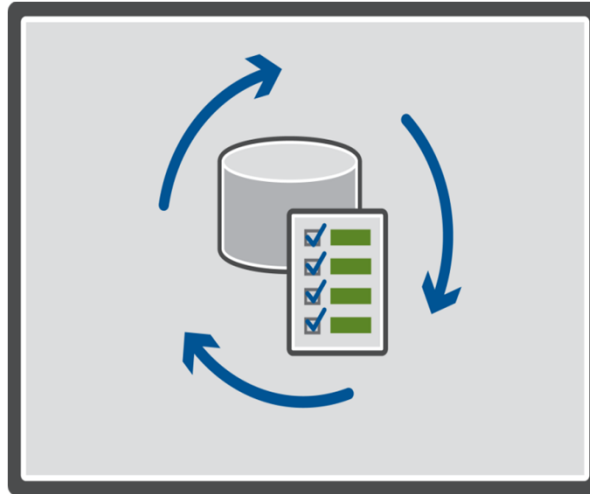
## Included in IDQ V9



OOTB Mappings,  
Mapplets, Rules



OOTB RTM  
Dictionaries



Data Quality 9

Fee-Based  
Content



Country Packages

## Annual Subscription

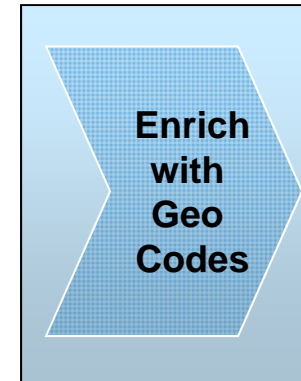
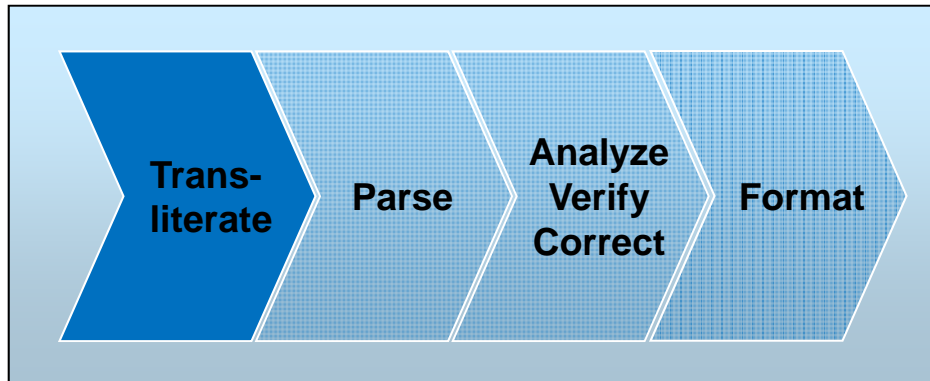


Address Validation  
Database  
Subscriptions



Geocoding Database  
Subscriptions

# Step 1 - Transliteration



Data input is in a non-Latin alphabet

ΑΘΗΝΑΣ 63  
105 52 ΑΘΗΝΑ  
GREECE

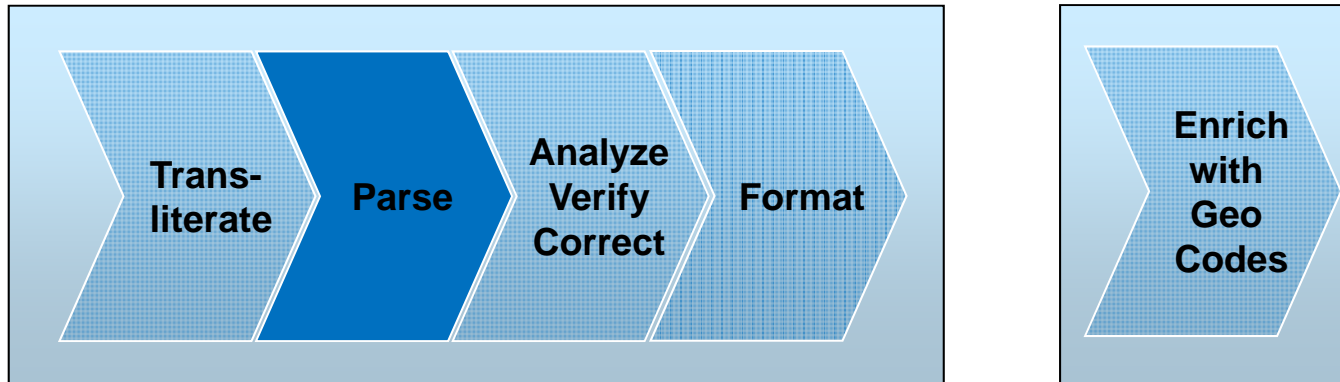


Data transliterated into Latin Script

ATHINAS 63  
105 52 ATHINA  
GREECE



# Step 2 – Parsing



Data is unstructured or in incorrect fields

7031 Columbia Gateway Dr  
Suite 101  
Columbia MD 21046  
USA

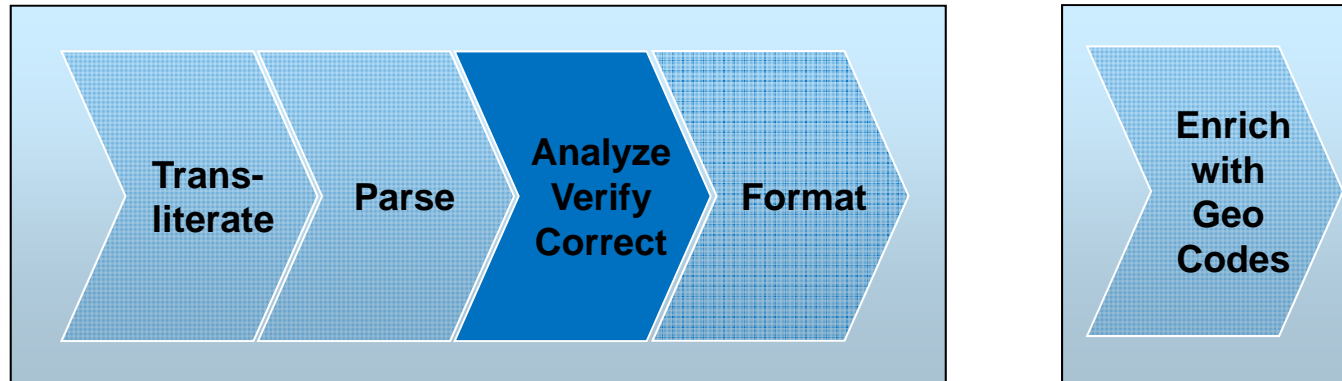


Parsing

All elements are stored in proper fields

House number: 7031  
Street: Columbia Gateway Dr  
Sub-Building: Suite 101  
City: Columbia  
State: MD  
ZIP: 21046  
Country: USA

# Step 3 – Correction



Incorrect input address

7031 **G**olumbia Gateway **Ave.**  
Suite 101  
Columbia MD 2104**4**  
USA

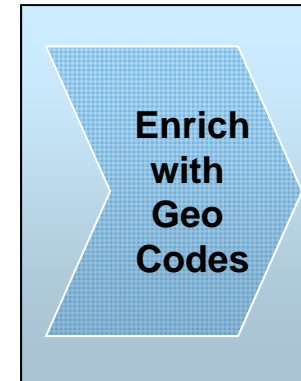
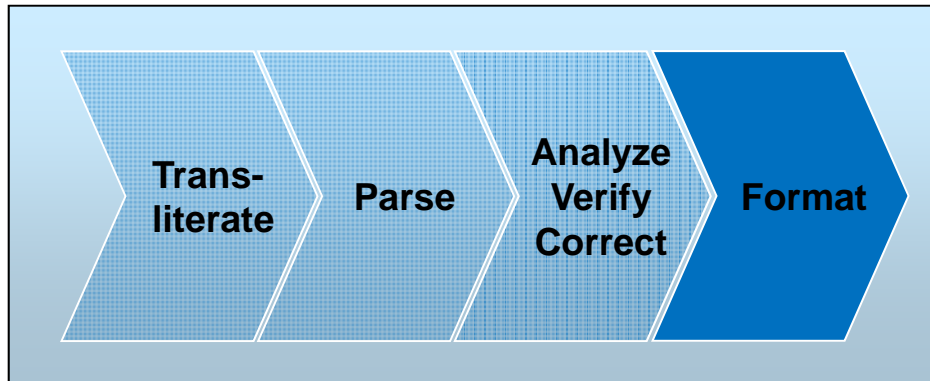


Correction

Corrected output address

7031 **C**OLUMBIA GATEWAY **DR**  
STE 101  
COLUMBIA MD 2104**6**  
USA

# Step 4 - Formatting



Properly formatted addresses for 240+ countries

Micros-Fidelio House  
6-8 The Grove  
Slough  
SL1 1QP  
Great Britain

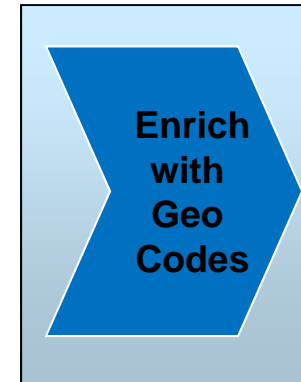
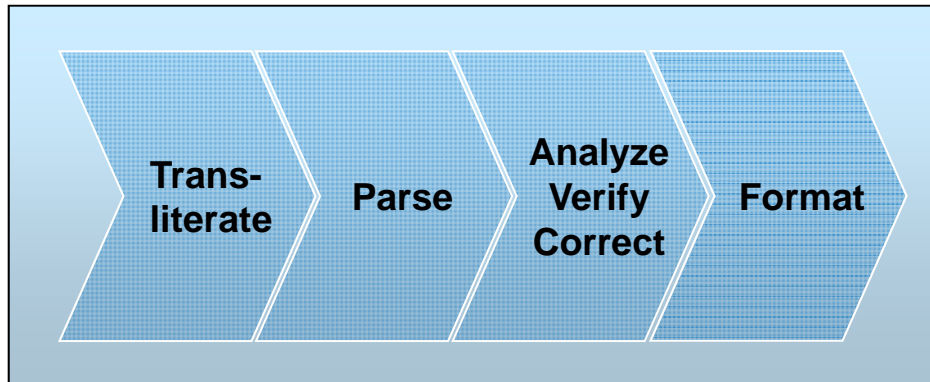
Europadamm 2-6  
41460 Neuss  
Germany

AΘHNAΣ 63  
105 52 AΘHNA  
GREECE

3440023 埼玉県春日部市大枝4-31-18  
Japan

7031 COLUMBIA GATEWAY DR  
STE 101  
COLUMBIA MD 21046  
USA

# Step 5 – Enrichment



Input address

7031 COLUMBIA GATEWAY DR  
STE 101  
COLUMBIA MD 21046  
USA



Geo Coordinates

Latitude: 39.173427  
Longitude: -76.801280

## Data Quality Matching



# Matching – Classic Data Matching

The Match transformation reads values in selected input columns and calculates match scores representing the degrees of similarity between the pairs of values.

## Match Strategies:

- Jaro Distance
  - Greatest emphasis on initial characters
- Bigram Distance
  - Longer multi token text strings
- Hamming Distance
  - Numeric, date & code
- Edit Distance
  - Strings of arbitrary length
- Reverse Hamming Distance



# Matching – Identity Matching

- **Identity Matching delivers next generation linguistic and statistical matching algorithms to ensure highly accurate matching out of the box for over 60 countries**
- **Identity Matching uses Populations contain key building algorithms that have been developed for specific countries and languages.**
  - Rules differ depending on the country/language

## Sample Rules

Category Name	Rule Type	Examples
Noise Word	Word is Deleted	e.g. THE , AND
Company Word Delete	Word is Deleted	e.g. INC , LTD , CO
Company Word Skip	Word is marked Skip	e.g. DEPARTMENT , ASSOCIATION
Personal Title Delete	Word is Deleted	e.g. MR , MRS , DR , JR
Nickname Replace Diminutives	Word and its Diminutives are Replaced	e.g. CATH(E , IE , Y) => CATHERINE
Nickname Replace	Word is Replaced	e.g. MIKE => MICHAEL
Secondary Lookup	Word generates additional search ranges	e.g. AL => ALBERT , ALFRED

# Classic Data Matching – product data

Type	Material	Shelf	Weight	Quantity	Packaging
DB	DB Biscs 9M 650g US	9M	0.676KG	20	Box
DB	DB Biscuits 9M 650g US	8M	0.667KG	20	Boxes
EDIT	BIGRAM	HAMMING	HAMMING	HAMMING	EDIT
1	0.878	0.5	0.6	1	0.6
Weights					
0.763					

Define the threshold that must be met before records will be output as a possible match



# Matching – Classic vs Identity

To the human eye certain names would be an obvious potential match.

However, classic match algorithms give poor results.

The Identity Population rules overcome this limitation without the need to cleanse or standardize data

## Classic Matching

	ClusterId	Name	LinkScore	ClusterSize	RowId	DriverId	LinkId
1	000000001	Kate Browne	1	4	1 - 1	1 - 4	1 - 1
2	000000001	Catherine Brown	0.5333333333...	4	1 - 2	1 - 4	1 - 1
3	000000001	Cathy Browne	0.75	4	1 - 3	1 - 4	1 - 1
4	000000001	Kathleen Brown	0.6428571428...	4	1 - 4	1 - 4	1 - 1

## Identity Matching

	ClusterId	Name	LinkScore	ClusterSize	RowId	DriverId	LinkId
1	000000001	Kate Browne	1	4	1 - 1	1 - 1	1 - 1
2	000000001	Catherine Brown	0.92	4	1 - 2	1 - 1	1 - 1
3	000000001	Cathy Browne	0.98	4	1 - 3	1 - 1	1 - 1
4	000000001	Kathleen Brown	0.96	4	1 - 4	1 - 1	1 - 2

# Sample matching output

The screenshot shows a data table with the following columns: ClusterId, GroupKey, ClusterSize, RowId, DriverId, DriverScore, LinkId, LinkScore, ID, COMPANY, and ADDR. Five callouts point to specific columns:

- Cluster ID:** Points to the ClusterId column.
- Cluster Size:** Points to the ClusterSize column.
- Sequence ID:** Points to the RowId column.
- Driver ID and Score: Record with the last ID in the Cluster:** Points to the DriverId and DriverScore columns.
- Link ID and Score: Record that brought the record in question into the Cluster:** Points to the LinkId and LinkScore columns.

ClusterId	GroupKey	ClusterSize	RowId	DriverId	DriverScore	LinkId	LinkScore	ID	COMPANY	ADDR
1	000000001	2	1 - 5336	1 - 5337	0	1 - 5336	1	5336	YEOVALE STOR...	1 MAI
2	000000001	2	1 - 5337	1 - 5337	0	1 - 5336	1	5337	YEOVALE STOR...	1 MAI
3	000000002	13	1 - 301	1 - 2239	0	1 - 301	1	301	WORLDSPAN	N8J1-
4	000000002	13	1 - 1868	1 - 2239	0	1 - 301	0.84723626765...	1868	WORLDSPAN	300 G
5	000000002	13	1 - 439	1 - 2239	0	1 - 301	0.87266666491...	439	WORLDSPAN	N8J1-
8	000000002	13	1 - 2239	1 - 2239	0	1 - 301	0.88266666491...	2239	WORLDSPAN	N8J1-
9	000000002	13	1 - 2183	1 - 2239	0	1 - 301	0.88266666491...	2183	WORLDSPAN	N8J1-
10	000000002	13	1 - 1981	1 - 2239	0	1 - 301	0.88266666491...	1981	WORLDSPAN	N8J1-
11	000000002	13	1 - 1980	1 - 2239	0	1 - 301	0.84723626765...	1980	WORLDSPAN	300 G
12	000000002	13	1 - 1691	1 - 2239	0	1 - 301	0.84861261176...	1691	WORLDSPAN	300 G
13	000000002	13	1 - 1755	1 - 2239	0	1 - 301	0.84723626765...	1755	WORLDSPAN	300 G
14	000000002	13	1 - 752	1 - 2239	0	1 - 301	0.87129032081...	752	WORLDSPAN	N8J1-
15	000000002	13	1 - 753	1 - 2239	0	1 - 301	0.86266666491...	753	WORLDSPAN	N8J1-
16	000000002	13	1 - 1103	1 - 2239	0	1 - 301	0.86266666491...	1103	WORLDSPAN	N8J1-
17	000000002	13	1 - 1320	1 - 2239	0	1 - 301	0.88266666491...	1320	WORLDSPAN	N8J1-
6	000000003	2	1 - 5490	1 - 5491	0	1 - 5490	1	5490	WHARFEDAILE...	458 S
7	000000003	2	1 - 5491	1 - 5491	0	1 - 5490	0.98399999976...	5491	WHARFEDAILE...	458 S
18	000000005	1	1 - 366	1 - 366	0	1 - 366	1	366	WINN DIXIE ST...	695 S
19	000000006	1	1 - 1102	1 - 1102	0	1 - 1102	1	1102	WINN DIXIE ST...	2080
20	000000007	1	1 - 116	1 - 116	0	1 - 116	1	116	WINN DIXIE ST...	805 S

# PowerCenter Integration



# Why Integrate with PowerCenter?

- **Deployment to PC for**
  - Performance
  - Reliability
    - PowerCenter HA
  - Scalability
    - PowerCenter Grid
  - Connectivity
  - Batch access
  - Web Services
  - DQ as part of ETL process

# IDQ/PC Integration Scenarios

## *New Install*

PC Version \ IDQ Version	9.1	9.0.1	9.0	8.6/8.6.1
9.1	No separate integration installer	DQ/PC Integration installers on PC side (Client and Server)	Not supported	8.6.1 only DQ/PC Integration installers on PC side (Client and Server)
9.0.1	Not supported	No separate integration installer.	Not supported	DQ/PC Integration installers on PC side (Client and Server)
9.0	Not supported	Not supported	No separate integration installer.	DQ/PC Integration installers on PC side (Client and Server)
8.6.2	Not supported	Not supported	Not supported	DQ/PC Integration installers for both IDQ and PC

# DQ v9.x in PC Upgrade

- **When upgrading only IDQ to 9.1:**
  - 9.1 PC Integration installers need to be run on PC side to allow DQ 9.1 mappings in PWC 9.0.1 / 8.6.1
- **When upgrading PC to 9.1:**
  - DQ 9.0/9.0.1 CTs will continue to run
  - Consolidation CTs will NOT be upgraded and will need to be re-exported from IDQ 9.1

# IDQ v8.6.2 in PC 8.6 -> 9.1 Upgrade

## There are 2 Options

### 1. Upgrade PowerCenter and apply CRS Upgrade patch.

- Run standard PowerCenter 8.6.x to PowerCenter 9.1 upgrade.
- Apply CRS Upgrade which is required to continue running mappings containing IDQ 8.6.x mapplets post upgrade to PowerCenter 9.x
- Requires 8.6.x Integration libs remain available.

### 2. Migrate IDQ Plans from 8.6.2 to 9.0.1

- Migration only from IDQ 8.6.2 to PowerCenter 9.0.1. No direct migration to 9.1
- Requires a number of steps
  - Export Data Quality 8.6.2 repository objects and reference data to the file system in XML format.
  - Convert the 8.6.2 repository objects to 9.0.1 format.
  - Import the reference data as reference tables to the 9.0.1 Model repository and staging area



**INFORMATICA**<sup>®</sup>  
The Data Integration Company<sup>™</sup>