



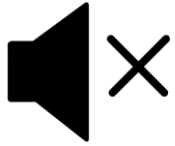
02 April, 2024

What's New in Informatica Cloud Data Governance and Catalog (CDGC), Cloud Data Marketplace (CDMP) and Cloud Data Quality (CDQ)? - April 2024 Release

- Amrish Thakkar, Senior Director, Product Management
- Kate Amory, Principal Product Manager, Product Management
- Markus Loechel, Senior Manager, Product Management
- Sean O'Shea, Senior Manager, Product Management

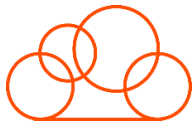
**Where data
& AI come to** 

Housekeeping Tips



- Today's Webinar is scheduled for **1 hour**
- The session will include a webcast and then your questions will be answered live at the end of the presentation
- All dial-in participants will be muted to enable the speakers to present without interruption
- Questions can be submitted to "All Panelists" via the **Q&A option** and we will respond at the end of the presentation
- The webinar is **being recorded** and will be available on our [Success Portal](#) - where you can download the **slide deck** for the presentation. The link to the recording will be emailed as well.
- Please take time to complete the **post-webinar survey** and provide your feedback and suggestions for upcoming topics.

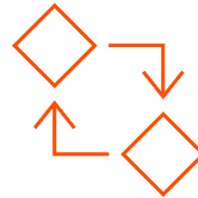
Feature Rich Success Portal



**Bootstrap trial and
POC Customers**



**Enriched Customer
Onboarding
experience**



**Product
Learning Paths
and Weekly
Expert Sessions**



**Informatica
Concierge**



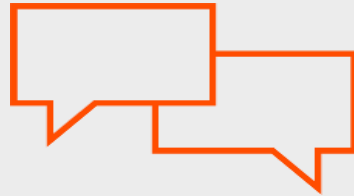
**Tailored training
and content
recommendations**

More Information



Success Portal

<https://success.informatica.com>



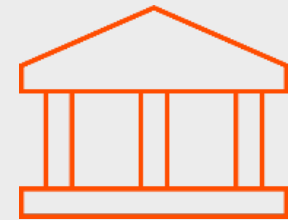
Communities & Support

<https://network.informatica.com>



Documentation

<https://docs.informatica.com>



University

<https://www.informatica.com/in/services-and-training/informatica-university.html>

Safe Harbor

The information being provided today is for informational purposes only. The development, release, and timing of any Informatica product or functionality described today remain at the sole discretion of Informatica and should not be relied upon in making a purchasing decision.

Statements made today are based on currently available information, which is subject to change. Such statements should not be relied upon as a representation, warranty or commitment to deliver specific products or functionality in the future.

Cloud Data Governance and Catalog

Features in CDGC 2024.04.M

- **Model and Relationship customization**
 - Including DQ Dimension, Data Classification Sensitivity
- **Updated Audit History**
- **Manual Data Element**
- **Obsolete Lifecycle for Business Assets**
- **Saved Searches**
- **CDGC to CDMP Integration**
- **Data Discovery Enhancements**
 - Limit Glossary Association to specific Glossaries hierarchy for a Catalog Source
 - Improved Discovery for Packaged applications (SAP, SFDC)
 - Increase Max Length of Classification Rules
 - Rename Data Element Classification
- **Scheduling Catalog Source Capability**
- **Lineage Enhancements**
 - Save Lineage Preferences
- **Usability Improvements**
 - Workflow requester
 - DQ Rule Occurrence for non-supported DE
 - CDQ Rule reference as part of Bulk Import
- **Language Support: Dutch for CDGC**
- **Azure Private Link support**
- **Scanner Additions and enhancements**
- **Profiling**

Model & Relationship Customization

Model Customisation

Attributes and Relationships now fully customisable

- Most OOTB attributes will now be customizable as custom attributes
- Customized in MCC as per screenshots
- Values can be added, renamed, deleted
- Among other things, you can customize
 - DQ Dimensions (max 12 values permitted)
 - Data Element sensitivity values (with new icon set)

The screenshot shows the 'dimension' configuration page in the Informatica Model Customization tool. The page includes a breadcrumb trail 'Dimensionize / Asset Customizations / All Assets / Attributes / dimension', a 'Save' button, and a 'Publish' button. The 'LIFECYCLE' section shows 'Draft' and 'Switch to Published'. The 'LAST UPDATED' section shows 'Jul 6, 2021, 22:05 PM' and 'Stephan Outry'. The 'Properties' section shows 'Name: Dimension', 'Data Type: Dropdown', and 'Subtype: Single Select'. The 'Values' section shows a table with columns 'Default' and 'Value'. The table contains the following rows:

| Default | Value |
|----------------------------------|--------------|
| <input checked="" type="radio"/> | Accuracy |
| <input type="radio"/> | Completeness |
| <input type="radio"/> | Consistency |
| <input type="radio"/> | Uniqueness |

The 'Consistency' row is highlighted. Below the table, there are icons for adding, editing, and deleting values. A 'Placeholder' label is visible at the bottom left of the table area.

A zoomed-in view of the 'Values' table is shown below the main screenshot, highlighting the 'Consistency' row and its associated icons.

| Default | Value |
|----------------------------------|------------|
| <input checked="" type="radio"/> | None |
| <input type="radio"/> | Low |
| <input type="radio"/> | Low- two |
| <input type="radio"/> | Low-Medium |
| <input type="radio"/> | Medium-one |
| <input type="radio"/> | Medium |

Relationship Types

- Extending range of relationship labels between existing assets
 - (doesn't allow creation of relationships between new asset pairs)
- All relationships between same asset types shown in one page
- Add / Remove
- Controlled – won't let users remove relationship types in use

The image shows two overlapping screenshots of the Informatica Metadata Command Center interface. The top screenshot displays the 'Business Term Relationship' configuration page. It includes a 'Properties' section with 'Source Asset Type' and 'Target Asset Type' both set to 'Business Term'. Below this is a 'Relationships' section with a 'Values' list containing 'Is classified by', 'Is described by', 'Is identified by', and 'Is related to'. The bottom screenshot shows the same configuration page with an 'Error' dialog box overlaid. The error message reads: 'You cannot Delete Relationship - Is generating. 55 relationships are impacted by the selected relationship value.' Below the error message is a table titled 'Asset Type (1)' with columns 'Asset Type' and 'Assets Impacted'. The table contains one row: 'Business Term' with '110' assets impacted. An 'OK' button is visible at the bottom right of the dialog.

Asset Audit History

Audit History

Reworked back end to provide reliable audit trail of asset activity

- All CDGC objects now have History tab
 - Captures all asset activity
- Filters for searching for specific actions
- Additions/deletions of child assets rolled up to the parent history
- Tickets also have updated look

| Date | Changed By | Event Type | Asset | Asset Type | Asset Path | Operation | From Asset Type | To Asset Type | Current Attributes | Previous Attributes |
|-----------------------|--------------------|---------------------|--------------|------------|--|-----------|-----------------|---------------|--|--|
| Feb 01, 2023 08:15 AM | Jerome Bell | Data Quality Result | NA_Customers | Table | | Created | | | Dimension : Validity Score : 88 Dimension : Uniqueness Score : 98 | |
| Feb 15, 2023 02:30 PM | Cameron Williamson | Lineage | NA_Customers | Table | | Created | Table | Table | From : GLOBAL_CUSTOMERS To : NA_CUSTOMERS | |
| Mar 05, 2023 09:45 AM | Esther Howard | Asset | CUST_NAME | Column | <input type="radio"/> NA_CUSTOMERS / <input type="radio"/> CUST_NAME | Deleted | | | Asset Type : Column Name : CUST_NAME | |
| Mar 19, 2023 09:00 PM | Jenny Wilson | Asset | CUST_SSN | Column | <input type="radio"/> NA_CUSTOMERS / <input type="radio"/> CUST_SSN | Created | | | Asset Type : Column Name : CUST_SSN Data Type : Varchar | |
| Apr 02, 2023 12:30 PM | Annette Black | Relationship | | | <input type="radio"/> NA_CUSTOMERS | Updated | Table | Business Term | Business Term : Customers Curation Status : Accepted | Business Term : Customers Curation Status : Recommended |
| Apr 16, 2023 04:45 PM | System (CLAIRE) | Glossary | NA_CUSTOMERS | Table | <input type="radio"/> NA_CUSTOMERS | Created | Table | Business Term | Business Term : Customers Confidence : 95% Curation Status : RECOMMENDED | |
| May 01, 2023 03:15 AM | Brooklyn Simmons | Stakeholder | NA_CUSTOMERS | Table | <input type="radio"/> NA_CUSTOMERS | Updated | Table | Stakeholder | User : User_007 Role : Data Steward, Data Owner | User : User_007 Role : Data Steward |
| May 15, 2023 06:00 PM | Courtney Henry | Stakeholder | NA_CUSTOMERS | Table | <input type="radio"/> NA_CUSTOMERS | Created | Table | Stakeholder | User : User_007 Role : Data Steward | |
| Jun 01, 2023 07:45 AM | Marvin McKinney | Asset | CUST_NAME | Column | <input type="radio"/> NA_CUSTOMERS / <input type="radio"/> CUST_NAME | Created | | | Asset Type : Column Name : CUST_Name Data Type : Varchar | |
| Jun 01, 2023 07:45 AM | Marvin McKinney | Asset | CUST_ID | Column | <input type="radio"/> NA_CUSTOMERS / <input type="radio"/> CUST_ID | Created | | | Asset Type : Column Name : CUST_ID Data Type : Varchar | |
| Jun 01, 2023 07:45 AM | Marvin McKinney | Asset | NA_CUSTOMERS | Table | <input type="radio"/> NA_CUSTOMERS | Created | | | Name : NA_CUSTOMERS | |

Manual Data Element

Manual Data Elements

New Business Asset for documenting assets unsuitable for technical catalog

- Business Data Sets only
 - Mix of technical and manual data elements in same business data set
- Allows business users to describe data they consume data for different use cases, e.g. where technical data is changed.
- Members of single business data set
- Relationships
 - Lineage /Glossary / Data Element Classification

The screenshot displays the Informatica Data Governance and Catalog interface. The main view shows a list of data elements for the 'Address Access DB' dataset. A context menu is open over the 'Manual Data Name' element, showing options to 'Add Data Element' and 'Create Manual Data Element'. The right sidebar provides details for the selected 'Manual Data Name' element, including its description, hierarchy, and associated glossaries.

| Name | Business Name | Type | Glossaries | Non-Distinct % | Source Data Type |
|------------------|---------------|---------------------|----------------------|----------------|------------------|
| AddressID | | Column | | | int |
| Manual Data Name | | Manual Data Element | Account Number... +3 | | int |
| AddressLine2 | | Column | Home address | | nvarchar |
| CUSTOMER_ID | | Access Table Column | | | |
| City | City | Column | City | World Cities | nvarchar |
| CUSTOMER_ID | | Access Table Column | | | |
| CUSTOMER_NAME | | Access Table Column | | | |
| ModifiedDate | | Column | | | datetime |
| PostalCode | Postal Code | Column | Postal Code | Postal_Code | nvarchar |
| rowguid | Postal Code | Column | Postal Code | Postal_Code | nvarchar |
| SpatialLocation | Postal Code | Column | Postal Code | Postal_Code | nvarchar |
| StateProvinceID | Postal Code | Column | Postal Code | Postal_Code | nvarchar |
| V_CUSTOMER_ID | Postal Code | Column | Postal Code | Postal_Code | nvarchar |

Lifecycle: Obsolete

New Lifecycle Value - Obsolete

Complete asset lifecycle journey, pre-deletion

- Allows upload of assets already in an Obsolete state – denotes they still have historical importance (e.g. future regulatory enquiry), not yet ready for deletion.
- May be particularly useful for customers migrating from Axon
- Implemented as fixed values, will be customisable in a later release

- ▼ Lifecycle (3)
- Published (37.6 L)
 - Draft (155)
 - In Review (8)
 - Obsolete (2)

Saved Searches

Saved Searches

- Users can save search queries with applied filters.
- Reusable Search definitions
 - Eliminates the need for users to repeatedly enter complex queries
- Users can add custom names to saved searches for easy retrieval.
- Saved Searches are private to the user who created them.

The screenshot displays the Informatika search interface. At the top left, a search bar contains the text "cdgc". To the right of the search bar is a "Save Search" button, which is highlighted with a red box. Below the search bar, the interface is divided into three main sections:

- Filters:** A sidebar on the left with a "Clear All" button. It shows "Asset Type (6)" with sub-filters for "Table (506)", "Business Term (3)", "Metric (2)", "Folder (2)", "Database (1)", and "Schema (1)". Below this is "Catalog Source Type (1)" with a "Select" dropdown.
- Search Results (506):** A central panel showing a list of results. The first result is "CDGC_COMPANY_INFO102", a TABLE located at "SEARCH_SQLSERVER_105K_CS / CDGC_VOLUME_TEST / dbo". The second result is "CDGC_COMPANY_INFO104", also a TABLE at the same location.
- Table Detail View:** A right-hand panel showing details for "CDGC_COMPANY... TABLE". It includes an "Overview" tab, "Related Assets (205)", a "Description" field, a "PUBLISHED" status indicator, "Data Collections" (with a note "Not added to Data Collections"), and a "Hierarchy" section showing the path "SEARCH_SQLS... / CDGC_VO... / dbo".

My Saved Searches

- Centralized storage for all saved searches
- Easy access and management of saved searches
- Mark search queries as favorites and sort based on favorites

New Search Based Widget

Create a widget based on a search query. If you run a particular search frequently, save the search as a widget. The widget always displays the up-to-date result of the search query.

Title: *

Description:

Select: *

Saved Search Type New Search Criteria

Search Criteria: *

Select Asset Types: *

Type: *

#

Preview Show Preview

Untitled

Dashboard Integration

- Ability to create dashboard widgets for quick access
- Customizable dashboard layout for personalized experience

My Saved Searches

9 Saved Searches

CDGC-60879_SAVED_SEARCH_1710511224609
CDGC-60879 Date Validation
FILTERS APPLIED: Created On: Last 365 Days Updated On: Last 200 Days

CDGC-60880_SAVED_SEARCH_1710516688467
CDGC-60880 Date Validation Grid View
FILTERS APPLIED: Created On: Last 70 Days Stakeholders: No stakeholders Updated On: From Jan 06, 2024 To Mar 14, 2024

More info ⓘ ☆

Edit Saved Search
Create Dashboard Widget
Delete Saved Search

CDGC to CDMP integration

Integration with Cloud Data Marketplace

- When viewing an asset that is linked to a Marketplace Data Product, you will see:
 - The teal badge indicating Data Marketplace presence
 - Links to the Data Products the asset resides within
- You will also see these in the preview panel from search results
- Clicking the 'info' tooltip next to a Data Product will open a preview showing key information about that Data Product

The screenshot displays the Informatica Data Governance and Catalog interface. The main view shows the asset 'V_FACT_ALL_SALES_ORDERS' with a teal 'Data Marketplace' badge. The interface includes a navigation sidebar, a search bar, and a main content area with tabs for Overview, Lineage, Relationships, Data Quality, Stakeholders, Attributes, and Tickets. A right-hand panel provides detailed information about the asset, including its description, stakeholders, data quality, and associated technical data sets. A bottom panel shows a preview for 'Anonymized Customer Data', which includes a Data Marketplace Rating of 4 stars, 76 consumers, 5 average days delivery, and managed access. The preview also shows a collection of vehicles leased by various organizations and a list of users who have requested, approved, or fulfilled the data product.

Informatica Data Governance and Catalog

Start typing search terms for suggestions

Sam Smith

NEW

Home

Search

Inbox

My Records

Dynamic Tab

John Jones

Overview

Relationships

Lineage

Hierarchical Struct...

snowflake_retal_deh / ... / EMEA_DEMO / CONTIGO /

V_FACT_ALL_SALES_ORDERS VIEW

SENSITIVITY LOW

LIFECYCLE PUBLISHED

DATA MARKETPLACE PUBLISHED

LAST UPDATED Jul 31, 2023, 1:38 AM

Overview Lineage Relationships Data Quality Stakeholders Attributes Tickets

Description

Attributes

Stakeholders

There are no stakeholders assigned to this asset.

Click here to add a stakeholder.

Data Quality

There are no data quality scores to show.

Associated Technical Data Sets

No data to display.

Status

Created By: Sam Smith

Updated By: Sam Smith

Data Marketplace Presence

Data Collection 1 Info

Data Collection 2 Info

Data Collection 3 Info

Data Collection 4 Info

V_FACT_ALL_SALES... VIEW

Overview Related Assets (13)

Description:

PUBLISHED

Related Data Collections:

Data Collection 1 Info

Data Collection 2 Info

★★★★☆ (1)

Hierarchy: snowflake_retal... / EMEA_DEMO... / CONT...

Data Quality: Rules Against Target

100%

↑+0%

Nov 27, 2023, 1:58 PM

Stakeholders:

JA John Admin Governance Admin...

MA Mark Adams Data Steward

Data Entity Classifications:

Associated Glossaries:

Anonymized Customer Data

Data Marketplace Rating: ★★★★★

Consumers: 76

Average Days Delivery: 5

Managed Access: Yes

Data Collection: Anonymized Customer Data

Category: Marketing

Certified Use: Marketing, Analytics

A collection of vehicles leased by:

- a business,
- government agency,
- or other organisation.

In the database, we keep records of how the fleet is organised, maintained and calibrated.

The Fleet Manager is responsible for planning, directing, managing, coordinating and supervising the programs for acquisition, assignment, utilization, maintenance, repair, replacement and disposal of fleet vehicles.

In the database, we keep records of how the fleet is organised, maintained and calibrated

What is it used for?

- Requested (3)
- Approved (3)
- Fulfilled (3)

Where is it used?

- Tableau Server (3)
- Snowflake Dat... (3)
- ADLS Gen2 (3)

1/2

Glossary Association & Data Classification

Glossary Association Scope

- In Glossary association configuration for a catalog source, you can now limit the scope of association to a specific hierarchy or hierarchies.
- Provides option to consider **entire glossary** or **subset of glossary**.
 - Allows you to select top-level glossaries
 - Only the business terms or metrics in selected hierarchies will be considered for glossary association
- Benefits :
 - More accurate associations
 - Better performance

Improve classification for packaged application (SAP, SFDC)

- In packaged applications such as SAP and Salesforce metadata objects names do not provide business context.
- To improve the overall classification of such sources, business names are extracted as part of metadata extraction and in generated classification instead of object names business names are considered for CLAIRE classification generation.
- Benefits :
 - This avoids invalid classification names.
 - Classification tags are more accurate

Increased max length of classification rules

- In Data Element Classifications length of the user classification rule is increased from 512 characters to 5000 characters.
- User can now add more conditions to fine tune classification rules.
- User can also validate rule length in MCC rule creation UI.

Rename Data Classifications

- User can now rename data element classifications and data entity classifications rules in MCC.
- Allows user to modify rule names for OOTB rules or user defined rules.
- Allows user to use rule names in local languages.
- Avoids unnecessary rule cloning and re-curation activity.

Other Enhancements

Schedule Catalog Source Capabilities

- In Catalog Source Configuration, user can now **add multiple schedules** in same catalog source configuration.
- User can schedule each capability or selected capabilities at different execution frequency or frequencies.
- Allows governance administrator to better manage capability dependencies.
- With the ability to schedule less source-intensive activities such as metadata extraction more frequently, administrators can ensure that essential metadata is always up-to-date without consuming excessive resources.
- Administrators can plan source-intensive jobs like profiling or data quality checks at different times based on factors such as data availability, system load, or business requirements. This allows for better optimization of system resources and ensures that these tasks are executed when they will have the least impact on other processes.

Save Lineage Preference

- Users can tailor their lineage viewing experience by saving preferences such as overlays, loops, or display link options.
- This customization ensures that users can focus on the specific aspects of lineage that are most relevant to their tasks or interests.
- Preference can be saved for technical lineage, business lineage or data element level lineages.
- User can reset saved preference to default at any time.
- User choice are preserved for lineage of all the assets in the catalog.
 - User will see same saved overlay options by default for any asset.

Data Quality Enhancement

- Allow association of Data Element with Rule Occurrence for Non DQ enabled sources
 - This allows for users to upload scores manually for the Catalog Source where DQ is not enabled
- Score propagation job to execute whenever a new score is pushed into Rule Occurrence via Bulk Import to give you an updated view quickly
- Ability to upload the Technical Rule Ref during Rule Occurrence bulk upload if autocatalog is switched on in MCC.
- Ability to Override Rule Occurrence schedule with the MCC Schedule defined for DQ Capability

Workflow Enhancement

- Ability to add Ticket creator / Ticket requestor as the task owner in workflow configuration
- Involvement of the requestor creates a feedback loop within the approval process enabling them to address any issues or questions raised by the approvers promptly.

Profiling Updates

- **PAM Additions**
 - Hive
 - PostgreSQL
 - Microsoft Fabric Warehouse
- **Random Sampling for AWS Redshift, Oracle and MS SQL Server**

Scanners Updates

New Scanners

- Microsoft Fabric OneLake (Preview)
- Microsoft Fabric Lakehouse (Preview)
- DBT (Preview)
- SAP S/4HANA Cloud (Preview)
- IBM DataStage

Scanners Updates 2024.04

Scanner Enhancements

- **Microsoft SQL Server Analysis Services enhancements**
 - Direct repository mode extraction (Tabular & Multidimensional)
- **Enable connection-aware scans for**
 - Microsoft PowerBI
 - Microsoft SQL Server Reporting Services
 - Tableau
- **Salesforce**
 - Support for Master-Details relationships
 - Connection assignment support
- **Databricks enhancements**
 - Databricks Workflows support
- **IDMC Metadata (Auto Cataloging)**
 - Support for CDI parameter files
- **Filter enhancements in:**
 - Lake family scanners
 - Redshift
 - MySQL
 - PostgreSQL
- **Authentication enhancements in ADLS Gen 2**
 - Shared Key Authentication
 - Managed Identity Authentication
- **Microsoft SQL Server linked servers support**
- **Oracle DBLinks support**
- **External Tables (Poly Base tables) support**
 - Microsoft SQL Server
 - Microsoft Synapse

Summary of Features in CDGC 2024.02.S

- **Browse Improvements**

- Find & Filter options
- Row count to indicate Children
- Browse option for Addl Asset types

- **Lineage Improvements**

- Tooltips to show addl details,
- Other usability enhancements

- **Improved Asset picker**

- **Quick Look enhancements**

- Support for SAP sources
- Support for Additional Technical Assets

- **Language Support**

- Italian in CDGC
- Japanese for QuickLook

- **Discovery Improvements**

- Allow for both Auto Generated and OOTB/User Defined classifications
- Ignore Prefix & Suffix in Glossary Association
- Bulk Promotion of Auto Generated Classification

- **Discovery (IGA, Data Classification, Relationship Discovery) PAM:**

- IICS, Hive, OneDrive

- **Profiling PAM**

- Teradata, SAP HANA, SAP BW, SAP BW in Hana, External Tables (Oracle, Azure Synapse, Athena)

- **Scanners**

- MariaDB, DB2 z/OS, Microsoft Fabric Data Warehouse

Cloud Data Marketplace

Where data & AI come to **LIFE**

Data Marketplace

- Model Alignment with CDGC – removal of 'Linked Assets'
 - Asset page is a window to CDGC - no local names, editability or comments
 - No more 'onboarding' CDGC Assets as linked assets
- Enhanced experience for adding assets to data products
- Usability Enhancements
 - Shield Badge for CDAM
 - Cost Centre Description displayed
- Additional language support
 - Italian
 - Dutch

The screenshot displays the Informatica Data Marketplace interface for a data product named 'Fleet DOE'. The interface is organized into several sections:

- Navigation:** A sidebar on the left contains options like 'New', 'Home', 'Browse', 'Search', 'Compare', 'My Data', 'Tasks', 'History', and 'Setup'. The 'Fleet DOE' product is selected in the sidebar.
- Product Summary:** At the top right, the product name 'Fleet DOE' is shown with a rating of 4 stars, a 'Published' status, and creation/last updated dates of 30 Sep 2022. Below this are tabs for 'Summary', 'Data Assets', 'Data Quality', 'Delivery', 'Terms of Use', 'Consumers', 'Approve', 'Fulfill', and 'Withdraw'. The 'Data Assets' tab is active.
- Data Assets Table:** A table lists individual data assets with columns for Name, Description, Data Source, Descriptive Source, Type, Status, Created, Last Updated, Rating, Data Quality, and Average. Assets include 'Wales University', 'Sports Undergrad', 'Funding 2023', and 'Wales univ...'. Each row shows a rating (e.g., 33, 11, 6) and a data quality percentage (e.g., 75%, 66%, 100%).
- Filters:** A sidebar on the left allows filtering by 'Certified' and 'Asset Type (13)'. Asset types include Column (380), Field (147), Table (78), View (73), Data Element (68), Native Data Asset (20), Native Data Element (12), Calculation (21), and Tableau Worksheet (4).
- Add Data Asset:** A section on the right allows adding new assets, with a search bar and a 'Finished Adding' button. A table lists assets like 'Reportaza', 'Mente Report', 'Sales Report', etc., with their descriptions, data sources, URIs, and types.
- Product Details:** A modal window titled 'Fleet DOE Fleet Product' is overlaid on the bottom right. It contains a description: 'Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.' Below this, it shows 'CERTIFIED USAGE' with 'Marketing' (pink dot) and 'Analytics' (blue dot) selected. At the bottom, it displays 'ORDERS' (76) and 'DELIVERY' (310 fulfilled, 6 days).

Scanners

Data Observability

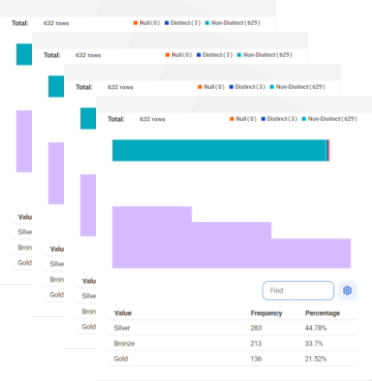
Data Observability

Overview

- New functionality to detect anomalies within the data automatically
 - Row count
 - Various profiling metrics such as number of distinct values, max length etc.
 - Changes in data types for data elements
- Enable and schedule for catalogue sources of your choice in the MCC
- Available for all sources that support profiling in MCC
- Adds valuable information about data anomalies to CDGC via new Data observability events

Overview

MCC Scan Job
(Data Profiling + DO enabled)



Results Comparison
& Event Generation

Visualize events within
CDGC

The screenshot shows a 'Data Observability' dashboard with the following components:

- DO Events (Metric):** A donut chart showing 12 events. Legend: Maxlength: 6 (orange), Minlength: 6 (purple).
- Maxlength change in Salary:** A 'DATA OBSERVABILITY EVENT' card with an 'Overview' button and a 'Summary' section. The summary text reads: 'The value usually remains constant but has suddenly changed.'
- Open Events by Metric:** A donut chart showing 2 events. Legend: Maxlength: 1 (orange), Minlength: 1 (purple).
- Open Events (2):** A table listing events:

| Name | Metric | Severity |
|----------------------------|-----------|----------|
| Minlength change in Salary | Minlength | Normal |
| Maxlength change in Salary | Maxlength | Normal |

The screenshot shows a data catalog entry for 'Salary' with a warning icon and the number '2'.

| | |
|----------------|---|
| CS_DQ_2 | 1 |
| sampleData.csv | 1 |
| Salary | 2 |

Data Observability

Configuration in the MCC

Configure and schedule

- New section in the source configuration “Data Profiling and Quality”
- Requires profiling without filters and in “full” scope

Informatica Metadata Command Center

+ New...
Home
Explore
Monitor
Configure
Customize
CS_DQ_1

• CS_DQ_1

Metadata Extraction | **Data Profiling and Quality** | Data Classification | Glossary Association

Data Profiling

Enable Data Profiling:

Connection and Runtime

Runtime Environment: invrh77DQCDev003.informatica.com

Filters

Specify additional filters on top of metadata filters.
 No Yes

Parameters

Modes of Run: Keep Signatures Only Keep Signatures and Values

Profiling Scope: Incremental Full

Maximum Precision of String Fields: 50

Date Format: MM/dd/yyyy HH:mm:ss

Data Quality

Data Observability

To get accurate results for data observability, set the Profiling Scope option to Full in the Data Profiling and Quality configuration.

Enable Data Observability:

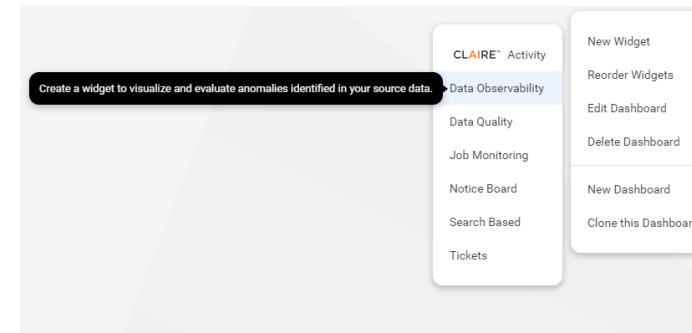
To get accurate results for data observability, set the Profiling Scope option to Full in the Data Profiling and Quality configuration.

Data Observability

Dashboard widget

Customized visibility

- New widget type “Data Observability”
- Select the sources that should be observed
- Donut and list type
- Show all open events for selected sources



New Data Observability Widget [Close]

Create a widget to visualize and evaluate anomalies identified in your source data.

Title: *
Observability events by metric

Description:
[Empty text area]

Catalog Sources: *
CS_DQ_1 x CS_DQ_2 x CS_DQ_3 x CS_DQ_4 x FS1 x [Dropdown]

Type: *
[Donut chart icon selected]

Group By: *
Metric x [Dropdown]

Preview [Refresh]

Observability events by metric

Group By: Metric

● Maxlength: 5
● Minlength: 5

[Help] [Create] [Cancel]

Data Observability

Technical asset view

Add visibility

- Overview with open events by severity
- New Data Observability tab
 - Donut charts grouped by metric and severity
 - List of events
 - Shows all events of this asset and child assets

The screenshot displays the Data Observability interface for a technical asset. The top navigation bar includes tabs for Overview, Contains, Lineage, Relationships, Data Quality, Data Observability (highlighted with an orange box), Stakeholders, Attributes, Tickets, and History. The main content area is divided into several sections:

- Description:** A text field for the asset's description.
- Glossaries:** A section showing accepted and declined glossary assets, with a recommendation for 'CLAIRE'.
- Data Entity Classifications:** A section showing accepted and declined classification assets.
- Data Elements (4):** A table listing data elements with columns for Name, Type, Glossaries, Classifications, and a Null/Distinct/Non-Distinct % bar.
- Open Events:** A donut chart showing 7 open events, all of Normal severity.
- Open Events by Metric:** A donut chart showing 7 open events, grouped by Maxlength (4) and Minlength (3).
- Open Events by Severity:** A donut chart showing 7 open events, all of Normal severity.
- Open Events (7):** A table listing the details of the 7 open events, including Name, Metric, Asset, Severity, Last Detected, and Status.

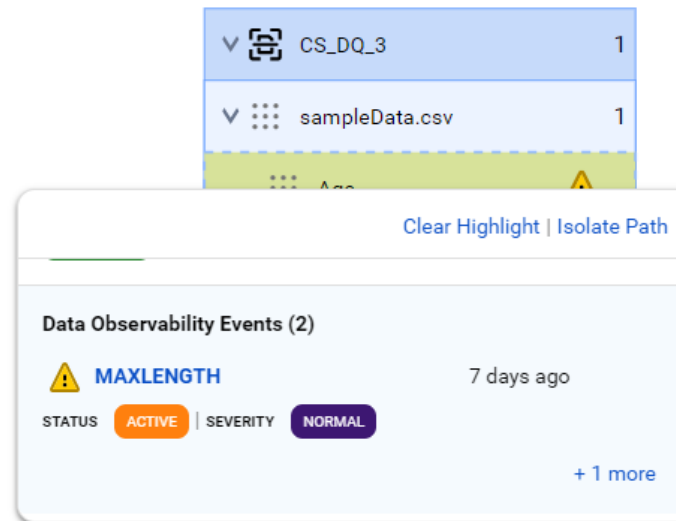
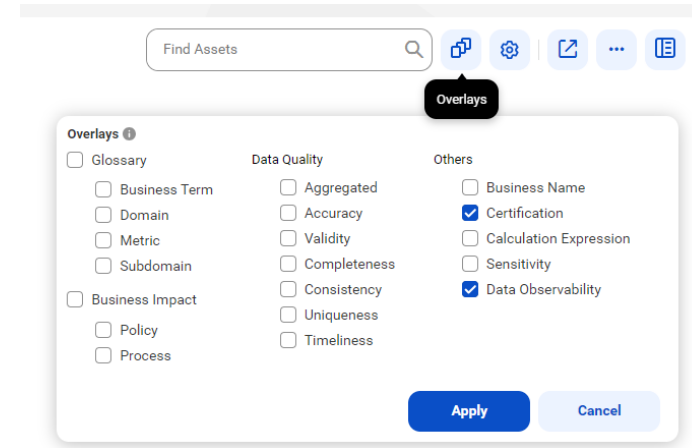
| Name | Metric | Asset | Severity | Last Detected | Status |
|----------------------------|-----------|--------|----------|---------------------------|--------|
| Maxlength change in Age | Maxlength | Age | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Minlength change in Height | Minlength | Height | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Maxlength change in Name | Maxlength | Name | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Maxlength change in Salary | Maxlength | Salary | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Minlength change in Salary | Minlength | Salary | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Minlength change in Age | Minlength | Age | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |
| Maxlength change in Height | Maxlength | Height | Normal | Mar 18, 2024, 09:55:42 AM | OPEN |

Data Observability

Lineage overlay

Add context and relations

- New overlay “Data Observability”
- Show data sets or elements with open events and the number of open events
- Preview for events

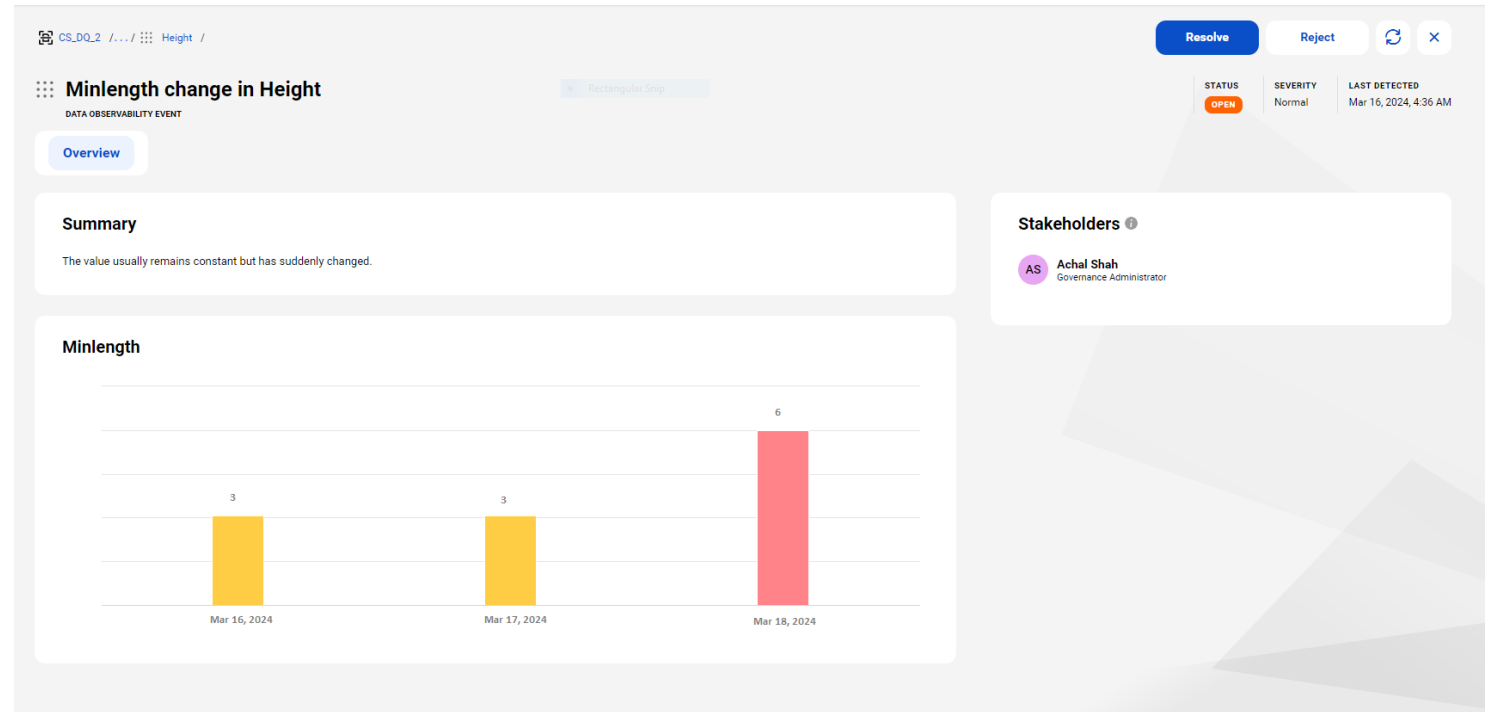


Data Observability

Event page

All information to take action

- Summary of what changed
- List of stakeholders inherited from the data set or data element
- Stakeholders can take actions
 - Resolve
 - Reject



Cloud Data Quality

Cloud Data Quality

What's New in Cloud Data Quality and Profiling

Data Quality

- **Data Observability**
- **New Visual Design**
- **CDQ available in OCI POD**
- New Data Quality Bundle for EU
- Exception Management Validation for Advanced Rules
- Run test on sub-org using parent org shared secure agent
- **Support for Azure Serverless (Native)**
- Platform Updates
 - MS Visual C upgrade to 2019
 - Rocky Linux 8 – replacement CentOS
 - Cortex XDR 8.2

Data Profiling

- **New Visual Design**
- **Product Availability Matrix (PAM)**
 - Profiling **Databricks** using **Secure Agent**
 - **Athena** connectivity using **EC2 Role**
 - **MariaDB** using **JDBC_V2**
- New public API for lookup profiling tasks
- Improvements to source mapplet support - SQL overwrite and bugs
- Platform updates
 - MS Visual C upgrade to 2019
 - Rocky Linux 8 – replacement CentOS
 - Cortex XDR 8.2



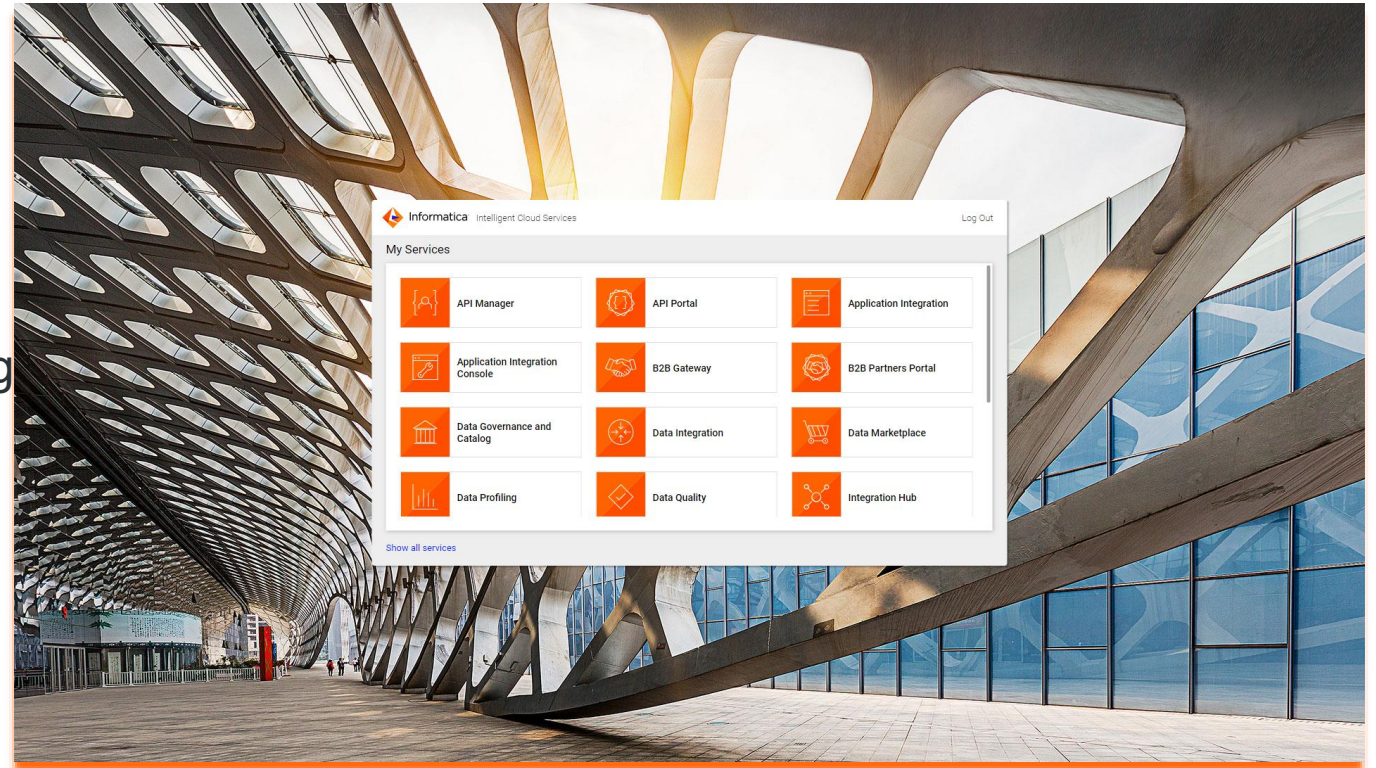
Designed for the business user

A New Visual Experience

Where data & AI come to **LIFE**

New Visual Design

- Objectives
 - Modernize the user interface
 - Improve UI response time
 - Align with other services already using the new design
 - Open doors for more future enhancements:
 - Accessibility
 - Dark Mode
 - Compact Mode



New UX Pattern Applied to CDQ Assets

New

Choose the asset type you want to create.

Rule Specification



Dictionary



Verifier



Cleanse



Deduplicate



Parse



Labeler



New Visual Design

Cloud Data Profiling

The screenshot displays the Informatica Data Profiling interface for a profile named 'Execute_snowflake_profile_native_98103'. The profile was run on March 18, 2024, at 6:07 PM, covering 107 rows across 7 columns and 0 rules.

The main view shows a table of column statistics:

| Columns | Value Distribution | % Null | # Null | % Distinct | # Distinct | % Non-di... | # Non-dis... | # Patterns | % of Top ... | Minimum... | Maxir |
|-----------|--------------------|--------|--------|------------|------------|-------------|--------------|------------|--------------|------------|-------|
| BOOLEAN | | 1.87% | 2 | 1.87% | 2 | 96.26% | 103 | 3 | 49.53% | 4 | 5 |
| INT | | 8.41% | 9 | 91.59% | 98 | 0% | 0 | 7 | 37.38% | 2 | 9 |
| DOUBLE | | 16.82% | 18 | 6.54% | 7 | 76.64% | 82 | 4 | 71.96% | 8 | 13 |
| FLOAT | | 49.53% | 53 | 50.47% | 54 | 0% | 0 | 4 | 49.53% | 3 | 4 |
| LONG | | 11.21% | 12 | 88.79% | 95 | 0% | 0 | 2 | 88.79% | 10 | 10 |
| TIMESTAMP | | 0.93% | 1 | 90.65% | 97 | 8.42% | 9 | 2 | 99.07% | 19 | 19 |
| STRING | | 11.21% | 12 | 0.93% | 1 | 87.86% | 94 | 2 | 88.79% | 253 | 253 |

Below the statistics table is a 'Data Preview (10 Rows)' section showing a sample of data:

| | BOOLEAN | INT | DOUBLE | FLOAT | LONG | TIMESTAMP | STRING |
|---|---------|-----------|---------------|-------|------------|-------------------------------|------------------------------|
| 1 | true | 12 | 9990000000000 | 7.16 | 4294967298 | 11/03/1998:00:00:00.000.+0530 | DdcJMJuwTsmgoygJjgsUZFBKR... |
| 2 | false | 123456789 | 9990000000000 | 8.16 | 4294967299 | 04/24/1998:00:00:00.000.+0530 | DdcJMJuwTsmgoygJjgsUZFBKR... |
| 3 | false | 123456790 | 9990000000000 | 9.16 | 4294967300 | 06/17/1987:00:00:00.000.+0530 | DdcJMJuwTsmgoygJjgsUZFBKR... |
| 4 | true | 123456791 | 9990000000000 | NULL | 4294967301 | 06/25/1997:00:00:00.000.+0530 | NULL |
| 5 | false | 123456792 | 9990000000000 | 5.12 | 4294967302 | 06/14/1996:00:00:00.000.+0530 | DdcJMJuwTsmgoygJjgsUZFBKR... |

The interface also includes a sidebar with navigation options like 'New...', 'Explore', and 'My Import/Export...'. The top right corner shows the Informatica logo and user profile icons.

New Visual Design

Cloud Data Profiling

Profile_TestException

Asset Details

Name: Profile_TestException

Description:

Location: Default

Created by: dq_reuser

Last Updated: Mar 18, 2024, 11:04 PM

Used Storage (MB): 0

Source Details

Connection: cleanse_source_connection_034

Source Object: DQ_Src_UK_Data.csv

CLARE capabilities: Enable

Profile Settings

Run profile on: All rows for complete analysis

Drill down: Yes No

Columns | Filters | Data Preview

Select columns to run the profile on. (Selected 0 of 11)

| Columns | Type | Native Data Type | Precision |
|--|----------|------------------|-----------|
| <input type="checkbox"/> SequenceId | UNSTRING | nstring | 255 |
| <input type="checkbox"/> GroupKey | UNSTRING | nstring | 255 |
| <input type="checkbox"/> Company | UNSTRING | nstring | 255 |
| <input type="checkbox"/> Address_Part1 | UNSTRING | nstring | 255 |
| <input type="checkbox"/> Address_Part2 | UNSTRING | nstring | 255 |

Profile_TestException - run - 1-1

Updated Mar 20, 2024, 11:27 AM

Job Properties

Instance Name: Profile_TestException - run - 1-1

Asset Name: Profile_TestException

Asset Type: Data Profiling Task

Started By: dq_reuser through UI

Start Time: Mar 20, 2024, 11:25 AM

End Time: Mar 20, 2024, 11:27 AM

Duration: 00:01:50

Runtime Environment: invawsepod102.informatica.com

Results

State: ✔ Success

Error Message: No errors encountered.

Profiled Rows: 65

Subtasks (4)

| Instance Name | Location | Start Time | End Time | Status |
|---|----------|------------------------|------------------------|--|
| Loading data from staging area to metric... | Default | Mar 20, 2024, 11:27 AM | Mar 20, 2024, 11:27 AM | ✔ Success |
| s_profiling_901_1_1-1 | Default | Mar 20, 2024, 11:26 AM | Mar 20, 2024, 11:27 AM | ✔ Success |
| Generating data profiling mappings-1 | Default | Mar 20, 2024, 11:25 AM | Mar 20, 2024, 11:25 AM | ✔ Success |
| Fetching the source row count-1 | Default | Mar 20, 2024, 11:25 AM | Mar 20, 2024, 11:25 AM | ✔ Success |

Profile_TestException

Profile run 1 of 1 | 0 of 11 Columns | 1 of 1 Rules | 65 Rows (All rows) | Mar 20, 2024, 11:25 AM

Results | Definition | Rules | Schedule | Insights

View: Columns and Rules | with: All Statistics

| Columns | Value Distribution | % Null | # Null | % Distinct | # Distinct | % Non-d... | # Non-d... | % Patterns | % of Top ... | Mini |
|----------------------|-------------------------|--------|--------|------------|------------|------------|------------|------------|--------------|------|
| RS_TestException | Input(s): Address_Part1 | | | | | | | | | |
| ExceptionDescription | | 0% | 0 | 1.54% | 1 | 98.46% | 64 | 1 | 100% | 0 |
| ExceptionPriority | | 0% | 0 | 1.54% | 1 | 98.46% | 64 | 1 | 100% | 0 |
| PrimaryRuleSet | | 0% | 0 | 1.54% | 1 | 98.46% | 64 | 1 | 100% | 0 |

Trend

| Type | Trend | % Change |
|--------------|-------|----------|
| Null | . | |
| Distinct | . | |
| Non-distinct | . | |

Data Types (2)

string(255) Documented Data Type

Data Preview (10 Rows)

| SequenceId | GroupKey | Company | Address_Part1 | Address_Part2 | Address_Part2_1 | Postal_Area | Person_Name | Title | Phone | Email | |
|------------|----------|--------------------|-----------------------|--------------------|-----------------|-----------------|-------------|--------------------|---------------------|-------------------|-----------|
| 1 | 199 | YAVALASTAR | Yeovale Stores | 1 Margrove Terrace | Barnstaple | Devon | EX32 7AD | MRS OPPORTUNA ... | Director of IS | Tel: 01271342156 | ifracomb |
| 2 | 200 | YAVALASTAR | Yeovale Stores | 1 Margrove Terrace | Barnstaple | Devon | EX32 7AD | MRS OPPORTUNA ... | Programmer | Tel: 01271342156 | ifracomb |
| 3 | 318 | WATRAS | Waitrose | 980 Oxford Road | Reading | Berkshire | RG30 6WR | Mr DERWON GRAS... | Owner/Consultant | Tel: 0118 9422444 | derwon_c |
| 4 | 473 | WASTSTRATGANANA... | West Street Genera... | 185 West St | Sheffield | South Yorkshire | S1 4EW | Mrs LACI COLOMB... | Computer Specialist | Tel: 0114249 1770 | sheffield |
| 5 | 474 | WASTSTRATGANANA... | West Street Genera... | 185 West St | Sheffield | South Yorkshire | S1 4EW | Mrs LACI COLOMB... | Owner/Consultant | Tel: 0114249 1770 | sheffield |

New Visual Design

Rule Specification

The screenshot shows the Informatica Data Quality interface for configuring a rule named 'rs_generic_valid_values'. The interface includes a left-hand navigation pane with a tree view containing items like 'New...', 'Home', 'Explore', 'My Jobs', 'My Import/Export...', and several rule objects. The main workspace is titled 'rs_generic_valid_values' and has tabs for 'Definition' and 'Configuration'. A 'Design' view is active, showing a table for rule logic definition. The table has columns for 'Inputs', 'Operator', 'Condition', and 'Action'. The rule logic is defined as follows:

| Inputs | Operator | Condition | Action |
|--------|-------------|----------------------------|---|
| if | InputValue | is not NULL | |
| and | ValidValues | is not NULL | then IIF(REG_MATCH(InputValue,ValidValues),'VALID','INVALID') |
| or | if | InputValue is NULL | then Valid |
| or | if | No rule statement is valid | then Invalid |

New Visual Design

Rule Specification

The screenshot shows the Informatica Data Quality interface. The top navigation bar includes the Informatica logo, "Data Quality", and user profile icons. The main content area is titled "rs_generic_valid_values_sean" with an "Advanced" badge. Below the title are "Save" and "Validate" buttons. The "Definition" tab is active, displaying the "Asset Details" section. This section includes a "Name" field with the value "rs_generic_valid_values_sean", a "Description" field with the text "Validates that the input value is one of the values checked in this rule specification. InputValue is the value being tested.", a "Location" field with the value "DQ_Shared/Rule_Specifications" and a "Browse" button, and a "Dimension" dropdown menu set to "None". A left sidebar contains a navigation menu with options like "New...", "Home", "Explore", "My Jobs", and "My Import/Export...".

The screenshot shows the Informatica Data Quality interface, specifically the "Rule Editor" for the rule "rs_generic_valid_values_sean". The top navigation bar is consistent with the previous screenshot. The "Configuration" tab is active. The "Rule Editor" section contains a code editor with the following SQL rule definition:

```
1 IF ValidValues != null AND InputValue != null THEN valid_values_check = IIF(REG_MATCH(InputValue,ValidValues),VALID,INVALID)
2 ELSEIF InputValue = null THEN valid_values_check = 'VALID'
3 ELSE valid_values_check = 'INVALID' ENDIF
4
5
```

Below the code editor is a "Test Data" section with a "Runtime Environment" dropdown menu set to "Choose an option:", a "Test" button, and a search bar. A left sidebar is also visible, showing a navigation menu.

Cloud Profiling

Profiling of Databricks

Secure agent – native mode

Use case

- Profiling of Databricks Delta Tables and External Tables

When to use Secure Agent – Native?

- When user requires data preview, drill down, filters, reduce the scope of profiled rows (First N), Data Quality Insights, SQL Override

CONS

- Cannot profile complex data types – content is treated as a string

When to use Advanced/Elastic?

- When user requires the profiling statistics for columns with **complex data types** (array, map, struct) is required

CONS

- Inability to profile rows (First N), drill down, add filter, data preview, SQL Override

[Detailed of features and functions supported in each runtime is here](#)

Cloud Data Profiling

What sources can you profile? (starting April 2024)

- Microsoft SQL and Azure Database with Active Directory Authentication
 - Microsoft Azure Synapse SQL (using Native driver)
 - Azure ADLS gen 2 and Azure Blob storage type
 - Oracle Database and Oracle Autonomous DB
 - Microsoft SQL and Azure Database
 - Amazon Redshift (using native V2 driver)
 - Amazon Athena
 - Snowflake (using native V2 driver)
 - Google Big Query
 - Delimited File on: File system | Amazon S3 | Azure ADLS Gen2 | Google Cloud Storage
 - ODBC
 - JDBC (NEW)
 - SAP ECC, SAP S/4 Hana and SAP BW
 - Databricks Delta tables using IDMC connector
 - Secure agent
 - Advanced / Elastic
 - Salesforce
 - Salesforce Sales Cloud | Salesforce Service Cloud
 - Salesforce Verticals:
 - Salesforce Financial Cloud
 - Salesforce Health Cloud
 - Applications on Force.com
 - Source mapplets
 - PARQUET, AVRO and JSON files on:
 - Amazon S3 - [Requires Advanced / Elastic]
 - Azure ADLS Gen2 - [Requires Advanced / Elastic]
- Deprecated**
- Azure Synapse, Snowflake, Redshift and Databricks using ODBC (still works, not recommended)
 - Databricks Delta tables using ODBC driver

Cloud Data Quality

Rule Specification Advanced

Adding Exception Management validation

- A Rule Specification can only be configured with Exception Management (EM) metadata at the top-level.
- This is consistent for both Basic and Advanced Modes.
- Validation added to support only one Rule Specification can be configured with EM metadata
 - If the script in advanced rule refers another advanced rule which is configured with exception management, then once user clicks on the validate button, the validation message gets displayed.

The screenshot displays the Informatica Rule Editor for 'RuleSpec2' in 'Advanced' mode. The interface includes a 'Save' button and a 'Validate' button. The 'Configuration' tab is active, and the 'Rule Editor' shows the following script:

```
1 IF rule:Default.RuleSpec1(Input).Output = "Eligible"  
2 THEN Output = "data processed"  
3 ELSE Output = "not processed"  
4 ENDIF
```

A 'Validation Errors' dialog box is overlaid on the editor, containing the following message:

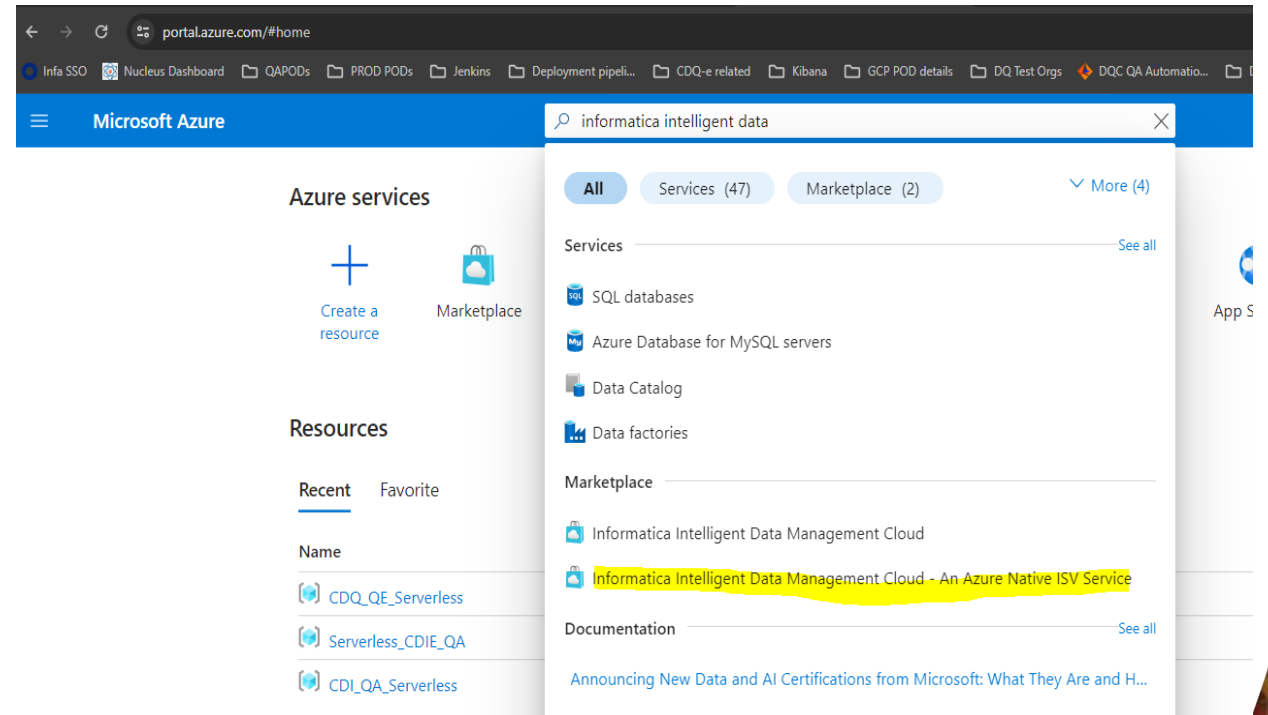
Validation Errors

A rule statement is not valid because it reads a rule specification that is configured for exception management. Select a different rule specification in the rule statement.

OK

Azure Serverless

- Azure Serverless is supported in Azure Pods only.
- User should have access to Azure Portal to setup Azure Serverless organization and Runtime.
- Create Informatica Organization from Azure Marketplace - **'Informatica Intelligent Data Management Cloud - An Azure Native ISV Service'**



Where data & AI come to



Where data & AI come to

