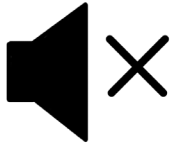


Apr 25th, 2023

What's New in Informatica Cloud Data Governance and Catalog (CDGC), Cloud Data Marketplace (CDMP) and Cloud Data Quality (CDQ)

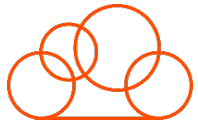
- Amrish Thakkar, Senior Director, Product Management
- Ian Stahl, Senior Director, Product Management
- Ricardo Sobral, Senior Director, Product Management

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 **INFASupport YouTube channel** and **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.

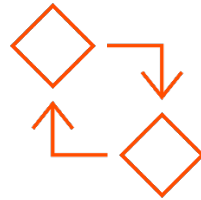
Feature Rich Success Portal



Bootstrap trial and
POC Customers



Enriched Customer
Onboarding
experience



Product Learning
Paths and Weekly
Expert Sessions



Informatica
Concierge



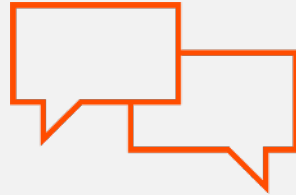
Tailored training and
content
recommendations

More Information



Success Portal

<https://success.informatica.com>



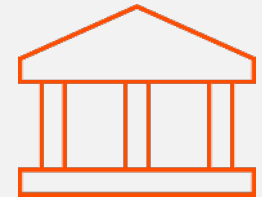
Communities & Support

<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 2023.04.M

- **Data Classification enhancements**

- Support for CLAI^RE generated Classifications
- Mark Classification as Sensitive
- Association of Governance Policies to data classifications

- **Dashboard enhancements**

- Search driven DQ Widgets
- Admin set Tenant level Default Dashboard

- **Auto-generation of Reference ID for Business Assets**

- **Custom Field enhancements**

- Support for Technical Assets
- Support for Multiple Asset Types
- Updated Preview Pane to show Custom Fields

- **Support for User Grid Preferences**

- Add/Remove columns, Rearrange and Resize Columns, Save user preference for Grid layout

- **Data Quality enhancements**

- Data Preview support for Data Quality rules
- Admin to be able to Turn On/OFF DQ Execution and Automation
- DQ Execution allowed on Sampled / Actual data
- Run Now option available for DQ Rule Execution

- **20+ New Scanners Added**

- **Profiling Updates**

- Databricks Tables, Custom Sampling support for Athena, Support for Profile Filters



Data Classification enhancements

CLAIRE Generated Classifications

Motivation

- Enrichment of data elements with classifications can be tedious
- Writing and evaluating Classification rules requires skill and time
- Many do not know what classifications should be even there
- Many OOTB classifications use data values while matching, but profiling is not allowed in many cases
- Many customers have well defined nomenclature of the asset names

Solution

- CLAIRE takes advantage of the well-defined nomenclature for asset names
- Generates classification labels for data elements which may be
 - Enabled for all catalog sources supported for classifications
 - Executed on a schedule or on demand for Catalog sources
 - Reviewed by users
 - Automated by users
- Easy discovery of Generated classifications through Semantic Search
- Profiling not needed

CLAIRE Generated Classifications – MCC

- Enable Generated classifications in MCC
- Monitor and understand what has been generated

✓ Metadata Extraction Data Profiling and Quality ✓ Data Classification Relationship Discovery Glossary Association Writeback

Enable Data Classification:

Use the selection option to choose between CLAIRE generated classifications and data classification rules.

Generated Data Classifications Data Classification Rules

powered by CLAIRE™

Add Data Classification

▼ Data Classification	Apr 13, 2023, 02:5...	Apr 13, 2023, 03:0...	Apr 13, 2023, 03:2...	00:14:23	COMPLETED
Auto Classification	Apr 13, 2023, 03:0...	Apr 13, 2023, 03:0...	Apr 13, 2023, 03:2...	00:14:23	COMPLETED

Auto Classification Details

Properties Results

1K
NUMBER OF INFERRED RELATIONSHIPS BETWEEN DATA ELEMENTS AND GENERATED CLASSIFICATIONS

1K
NUMBER OF PROMOTED RELATIONSHIPS BETWEEN DATA ELEMENT AND DATA ELEMENT CLASSIFICATIONS

Sensitivity Level with Data Element Classification

- Users can set a sensitivity level for a classification in Metadata Command Center (MCC)
- The sensitivity level can be set as the following
 - High
 - Medium
 - Low (no sensitivity)
- The OOTB classifications are pre-defined with a sensitivity level.
- On Explore, you can identify the sensitive data element cl

• Albania National ID For Demonstration

1 General Information 2 Qualifier

Name: * Albania National ID For Demonstration






Description: This matches with Albanian ID. This will verify the patterns of the ID like 930131123A or 910231123B or A00101123A.

Type: Element

Sensitivity

Sensitivity Level: * High

Data Classifications (338)

Name ↑	Type	Description	Sensitivity	Updated On
 Albania National ID	Data Element	This matches with ...	HIGH	Apr 11, 2023, 7:22 ...
 Amex	Data Element	This identifies Am...	HIGH	Apr 11, 2023, 7:20 ...
 Argentina National ID (DNI)	Data Element	This matches with ...	HIGH	Apr 11, 2023, 7:22 ...
 Australia Bank Account Number	Data Element	Classifies metadat...	HIGH	Apr 11, 2023, 10:0...
 Australia Passport	Data Element	Classifies metadat...	HIGH	Apr 11, 2023, 7:20 ...

Searching for sensitive data assets

- Find classifications with a sensitivity level
 - Query: *classifications with sensitivity in (high, medium)*
- Find assets with classifications with marked sensitivity level
 - Query: *data elements related to classifications with sensitivity in (high, medium)*
- The data element overview clearly indicates the sensitivity level

The screenshot shows a search interface with the following components:

- Filters:** Asset Type (1), Lifecycle (1) with 'PUBLISHED (3)' selected, and Created By (1).
- Search Results (3):** A list of three assets: BRNCH_CD, MANGO, and MANGO. Each asset entry includes its name, type (COLUMN), and associated classification codes (e.g., CDGC30143, EDC_HAWK, BIGDATASIZE).
- Asset Overview (BRNCH_CD):** Shows a 'Description' field, a 'LIFECYCLE' status of 'PUBLISHED' with a 'HIGH' sensitivity indicator (highlighted by a red arrow), and a 'Hierarchy' path.

The screenshot shows the detailed overview for the BRNCH_CD asset, including:

- Metadata:** Business Name (Add Business Name), Sensitivity (HIGH), Lifecycle (PUBLISHED), and Last Updated (Apr 11, 2023, 8:35 PM).
- Navigation Tabs:** Overview (selected), Lineage, Relationships, Data Quality, Stakeholders, Properties, Tickets, History.
- Description:** A text field for the asset's description.
- Stakeholders:** A section indicating 'There are no stakeholders assigned to this asset.'
- Catalog Source Definition:** A section for defining the asset's source.
- Data Quality:** A section indicating 'There are no data quality scores to show.'

Policy to Classification Relationship

- Users can manually associate Data Element and Entity classifications to defined Policies
- Helps in identifying the Assets indirectly linked to a Governance Policy via Data Element / Data Entity classification through Search queries.

The screenshot displays the Informatica Governance console interface. On the left, a sidebar shows the 'PII - Personally Identifiable Information' policy with tabs for Overview, Hierarchy, and Relationships. The main area is titled 'Add Relationships' and features a search bar with the query 'data element related to Policy PII*'. Below the search bar, two search result panels are visible, each showing two results for the 'SSN' column. The first result is from 'Snowflake_Profiling_P0' in the 'EDC_HAWK / EDC_HAWK_SCHEMA / EMPLOYEE_TEST' context. The second result is from 'SANITY_ORACLE_SOURCE_1680807449941' in the 'RND12CR2 / LINEAGE_VALIDATION / EMPLOYEES_DS1' context. On the right, a detailed view of the 'SSN' column is shown, indicating it is a 'COLUMN' with 5 related assets (2 views, 1 table, 1 data set) and 1 data element classification named 'Custom_SSN'.



Dashboard updates

Tenant specific Default Dashboard

Subtitle goes here if needed, Roboto 24 pt, one line

- Provides an ability for the Admin user to set a Default Dashboard for the Tenant
 - This allows for the OOTB dashboard to be replaced with Tenant specific dashboard
 - Users will land on this dashboard by default unless they have changed their default
 - This dashboard cannot be edited by the users. They can clone it if they want to create a copy and make changes to it.
- Only one dashboard can be shared currently and this becomes the default for all users
 - Users can change their default setting at any point

Search based DQ widget

Subtitle goes here if needed, Roboto 24 pt, one line

- Users can now create a DQ widget based on Search criteria
 - This allows the users to create multiple DQ widgets based on logical separation they want
- Search option is only allowed for “Donut chart” visualization
 - This is not available for “Bar chart” or “Score chart”



Data Quality related updates

Data Quality – Data Preview

- DQ Failed Rows Preview

- Ability to do Data Preview of Failed rows in CDGC App at the Rule Occurrence
- Data Preview only available for user with Privilege 'View Sensitive Data'
- Possible to see both Cache / Live Data – Configured at MCC

Preview of Failed Rows

RO_TestFloat_Snowflake_Ria(NUMBER_COL): 48 Rows

ROWID_COL	INTEGER_COL	NUMBER_COL	TIMESTAMP_COL	SMALLINT_COL	FLOAT_COL	NVARCHAR2_COL
NULL	80	0.913566881	NULL	NULL	2.40566e+58	Smith
AAAAooAATAAABr...	50	0.905981378	NULL	NULL	9.94684e+58	Sciarra
NULL	40	0.954563304	NULL	NULL	7.90178e+58	OConnell
NULL	80	NULL	NULL	NULL	1.97008e+58	NULL
NULL	80	NULL	NULL	NULL	6.54282e+58	NULL
NULL	30	NULL	NULL	NULL	5.32382e+58	NULL
NULL	80	NULL	NULL	NULL	5.44209e+58	NULL
NULL	30	NULL	NULL	NULL	6.45237e+58	NULL
NULL	50	NULL	NULL	NULL	9.44939e+58	NULL
NULL	50	NULL	NULL	NULL	2.33493e+58	NULL
NULL	80	NULL	NULL	8547	5.27981e+58	NULL
NULL	50	NULL	NULL	NULL	1.84057e+58	NULL

Cache Result: * ?

Sampling Type: * ?

Agent Cache

Agent Cache

No Cache

Data Quality – Run Now at Rule Occurrence

- On demand execution of Rule Occurrence via Run Now option
 - Run Now option only available for the stakeholder
 - Ad-Hoc option removed from Rule Template / Rule Occurrence

The screenshot displays the Informatica Data Quality console interface for a specific Rule Occurrence. The breadcrumb navigation at the top reads: Snowflake_CS_Ria / INFADQ_DB / DQ_SCHEMA / CP_ALL_SUPPORTED_DATATYPE / NUMBER_COL. The main title is 'RO_TestFloat_Snowflake_Ria' with the subtitle 'DATA QUALITY RULE OCCURRENCE'. A 'Run Now' button is prominently displayed in the top right corner, alongside other action icons. Below the title, there are tabs for 'Overview', 'Score', 'Stakeholders', 'Tickets', and 'History'. The 'Overview' tab is active, showing a 'Description' field with the value 'RO_TestFloat_Snowflake_Ria'. To the right, there are three summary cards: 'OVERALL RATING' (5 stars), 'LIFECYCLE' (PUBLISHED), and 'LAST UPDATED' (Apr 5, 2023, 11:26 PM). Below the description, there are two columns of metadata: 'Rule Template:' (value: -) and 'Reference ID:' (value: RF_004) on the left; 'Sync with Rule Template:' (value: -) and 'Primary Data Element:' (value: NUMBER_COL) on the right. On the right side of the console, there are two panels: 'Stakeholders' showing 'CDGC QA' (Governance Administrator) and a 'View All' link; and 'Status' showing 'Created By:' 'QA QA on Apr 5, 2023, 7:31 PM'.

Data Quality – Execution On/Off

- DQ Capability at the Catalog source
 - Can be switched On / Off at the source
 - Ability to execute Data Quality check at the Sync Run in MCC
 - Ability to execute Data Quality Checks on Sampled Data / Full Data

The screenshot shows the configuration page for a catalog source named '100T_81C_1KR_Oracle'. At the top, there are navigation buttons: '< Back', 'Next >', and 'Save'. Below this is a horizontal menu with tabs: 'Metadata Extraction', 'Data Profiling and Quality' (which is selected), 'Data Classification', 'Relationship Discovery', and 'Glossary Association'. The main content area is divided into sections: 'Data Profiling' and 'Data Quality'. In the 'Data Quality' section, there is a toggle switch for 'Enable Data Quality' which is currently turned on. Below the toggle is a button labeled 'Use Data Profiling Parameters'. At the bottom of the page, there is a 'Connection and Runtime' section with a label 'Runtime Environment:' and a dropdown menu that currently shows 'Select a runtime environment'.

The screenshot shows a modal dialog box titled 'Run Catalog Source Job: 100T_81C_1KR_Oracle'. The dialog has a close button (X) in the top right corner. Under the heading 'Scope of Run', there are three checked items: 'Metadata Extraction', 'Data Profiling and Quality', and 'Data Quality'. At the bottom of the dialog, there is a question mark icon on the left, a blue 'Run' button in the center, and a light blue 'Cancel' button on the right.

Data Quality – DQ Automation

- DQ Automation Control at MCC
 - DQ Automation can be switched On/Off at the source level . If switched OFF only manual rule can be created for those sources
 - Possible to apply DQ Automation at all DE / Only DE linked to Business Dataset

Parameters

Data Quality Automation: * ?

Yes No

Cache Result: * ?

Apply on Data Elements linked with Business Dataset

Apply on all Data Elements

Sampling Type: * ?

Limit N Rows

Other improvements

Auto Generation of Ref ID

- System will now Auto Generate Ref ID for Governance Assets when left empty
 - This applies to both assets created from UI as well as Bulk Import
 - It will use the following format for the generated Ref IDs
<Prefix>-<Number>
- Once the Number has been used, it will 'not' be used again for Auto Generation.
 - For e.g. if the asset with Ref ID 'BT-1' gets deleted, system will not use 'BT-1' for Auto generation
- System continues to allow user to provide Ref ID manually as it was prior to this release.
- Prefix for each Asset Type can be configured in MCC

Custom Field enhancements

- Ability to assign field to multiple asset types
- Visualize custom field values in preview panels
- Ability to create fields for technical assets
 - Creation of custom field is per catalog source type
 - Asset type can be filtered by Core types (Dataset, data element, data process, etc...)
 - Once created, custom attributes are populated with default values or through CDGC Application

The screenshot displays the 'EMPLOYEES_DS1' table configuration page in the Informatica Data Catalog. The interface is divided into several sections:

- Assets:** Shows the asset path: CDGC_ORACLE_SCAN_001 / RND12CR2 / EMPLOYEES_DS1.
- Description:** Includes a 'PUBLISHED' status and a 'Hierarchy' path: CDGC_ORACLE_SCAN_001 / RND12CR2 / LINEAGE_VALIDATION.
- Custom Fields:** A red box highlights this section, showing a custom field named 'CDGC_TEST_DROPDOWN_TEC...' with a value of 'Improvement in Data Quality Sc...' and a code 'CDGC_002'.
- Data Quality:** Currently empty.
- Stakeholders:** Currently empty.
- Related Assets:** A diagram shows a 'View' (2) connected to the 'EMPLOYEES_DS1' table, which is connected to a 'Column' (9).

Grid Customization

Subtitle goes here if needed, Roboto 24 pt, one line

- Users can customize the Grids and perform following Actions:
 - Add/Remove Columns, Rearrange Columns, Resize Columns, Density/Mode
 - This settings are available for the following grids:

Asset Page : Contains Tab (where applicable), Relationship Tab, Tickets Tab, History Tab,

Browse

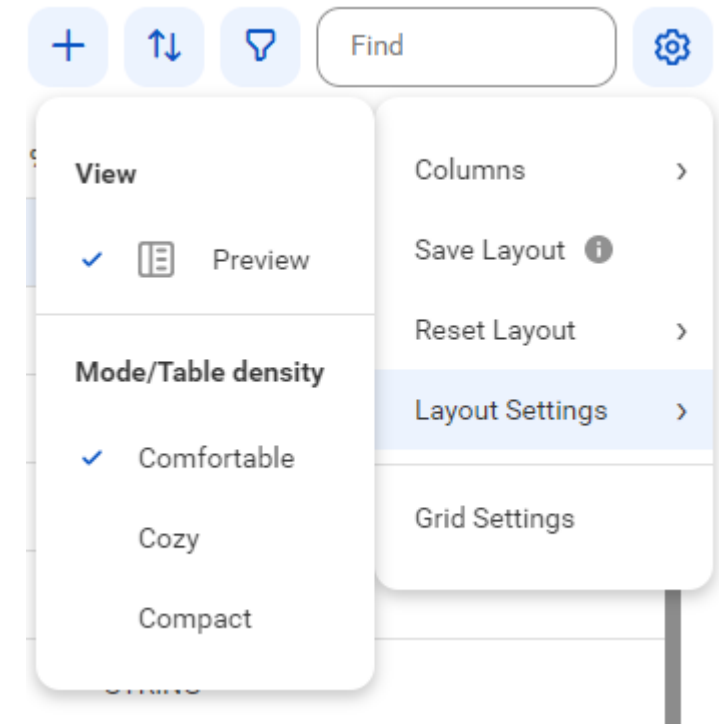
Search

Tickets manager List Page

Claire Insight table(Recommendations table) - Claire Dashboard

Job Monitoring Page

- User can reset to the Default or last view at any point



Key Profiling Enhancements

- Profiling enabled for Databricks delta tables
- Profiling filter supported
- Custom Sampling support for Athena



Scanners

Scanner Updates (Databricks – Unity Catalog Support)

- You can use the Databricks catalog source to extract metadata from all available catalogs including legacy Hive Metastore (Delta Lake) and Unity Catalog

Include Metadata Catalog Schema Pattern

Include Metadata Notebooks Path Pattern

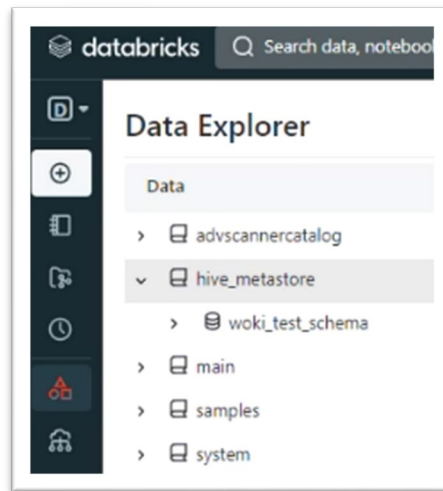
Configuration Parameters Show Advanced

Databricks Notebooks Workspace Path:

Databricks Notebooks Python Modules Path:

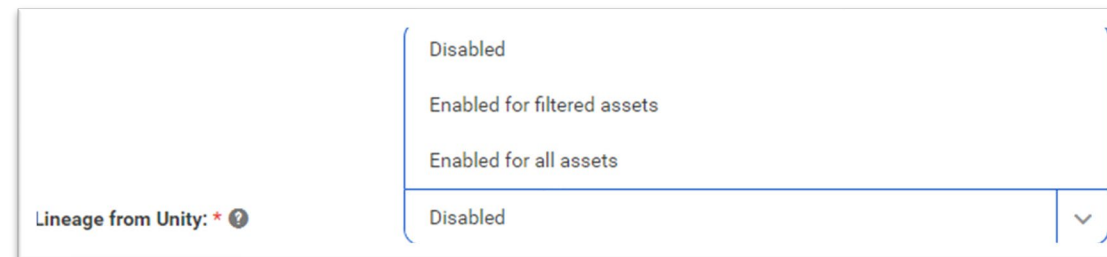
Lineage from Unity:

Scanner Updates (Databricks – Unity Catalog Support)



Name	Type
> dbfs	File System
> hive_metastore	Database
▼ samples	Database
default	Schema
nyctaxi	Schema
tpch	Schema



- You can extract lineage available in Databricks Unity Catalog (complete or partial, related to the Unity Catalog assets selected in the filter)







Scanner Updates (Database Scripts Scanners)

Enables metadata extraction from scripts storing SQL statements in various flavors


Specify metadata filters.


No filters  Filter conditions 

> Show supported wildcard characters and examples

Include  Script Path  Inner_Join_Insert.sql  

Configuration Parameters [Show Advanced](#)

Scripts Root Directory Path: * 

Default Schema: * 

Predefined Catalog Sources (5)

- IBM DB2 LUW Script**
A set of IBM DB2 SQL statements stored in a file which you can use for sequential execution.
- Microsoft SQL Server Script**
A set of Microsoft SQL statements stored in a file which you can use for sequential execution.
- Oracle PL/SQL Script**
A set of Oracle SQL statements stored in a file which you can use for sequential execution.
- Snowflake SQL Script**
A set of Snowflake SQL statements stored in a file which you can use for sequential execution.
- Teradata BTEQ Script**
A set of BTEQ commands and Teradata SQL statements stored in files that you can use for sequential execution.

Scanner Updates (Microsoft SQL Server)

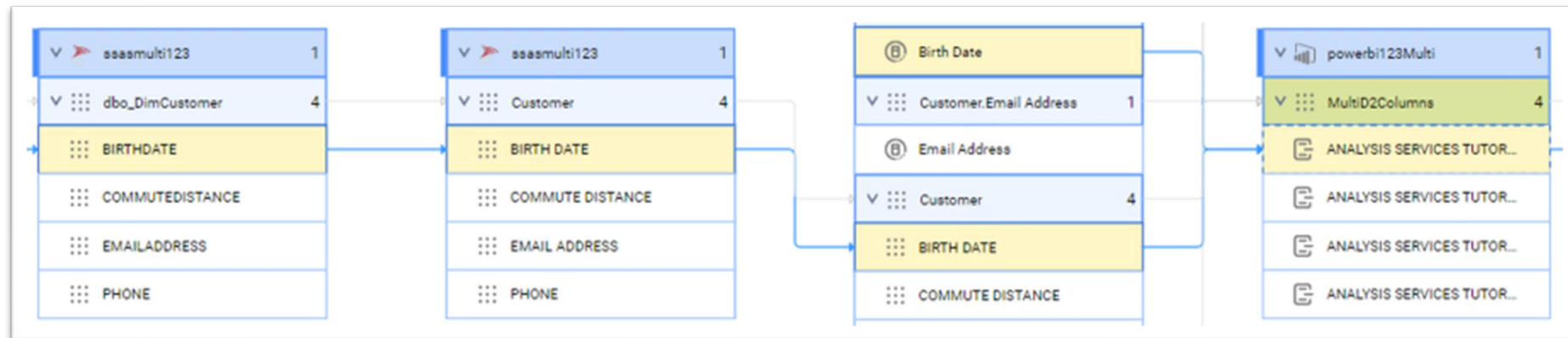
Microsoft SQL Server Analysis Services

- Both models are supported
 - Multidimensional
 - Tabular
- SSAS assets as a source of PowerBI reports



Microsoft SQL Server Analysis Services

An online analytical processing and data mining tool.

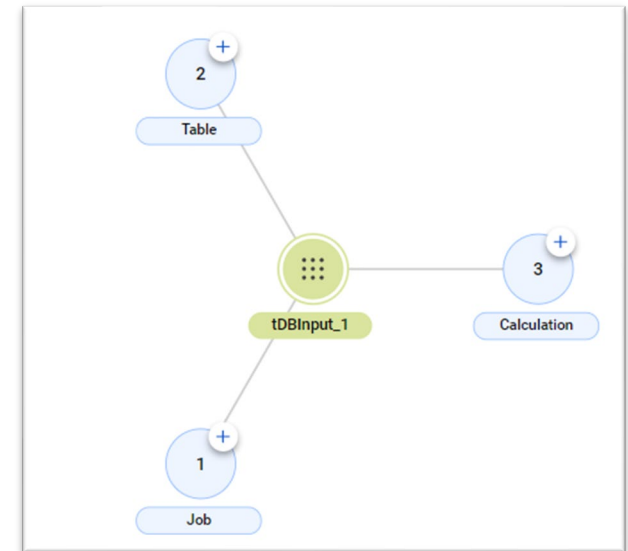
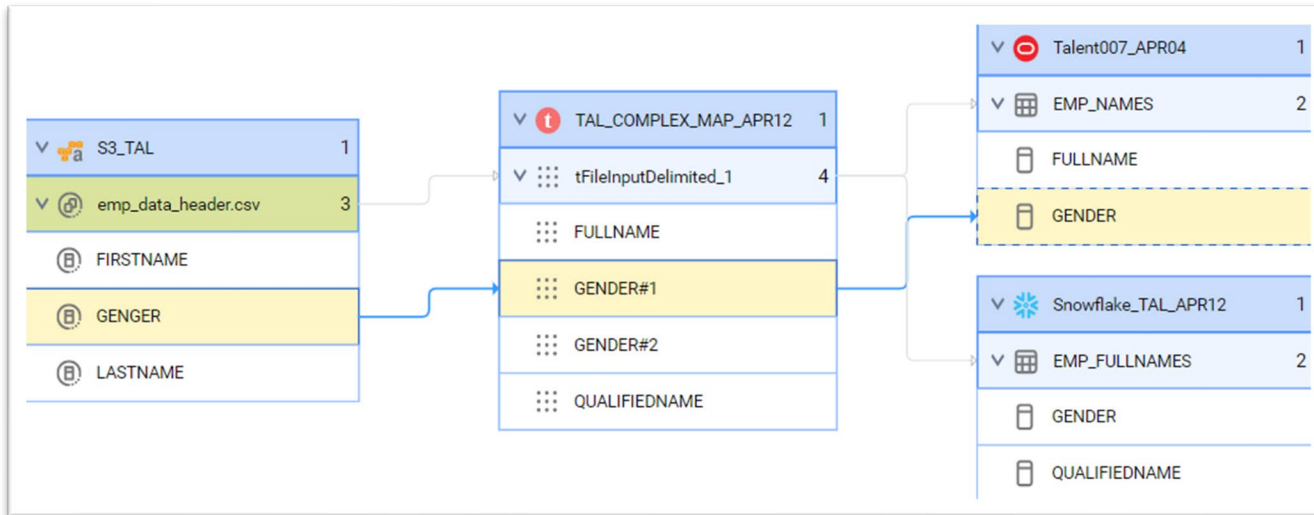


Scanner Updates (Talend DI)

Talend Data integration

You can extract metadata from the Talend Jobs specified as a scope in the scanner filter.

Talend Data Integration
An ETL data integration platform that you can use to gain business insights from data.



Scanner Update (File Systems)

Various new “file systems”

You can extract information about the individual files and their location from:

- Oracle Cloud Object Storage
- Microsoft OneDrive
- Microsoft SharePoint
- SFTP

The screenshot shows a sidebar with a list of file systems and a main panel displaying details for the selected 'File System (1)'. The sidebar list includes:

File System	Count
File System	1
Flat Field	2
Flat File	1
Folder	3
Hierarchical F...	47
Hierarchical F...	4

The main panel shows details for 'File System (1)':

Name ↑
Onedrive

The screenshot shows the 'OCI_IncludeCSVFile' source interface. The 'Overview' tab is selected, showing a summary of 'Flat Field (20)' and a list of individual fields.

Summary:

Flat Field	Count
Flat File	1
Folder	2
OCI Bucket	1

Fields List:

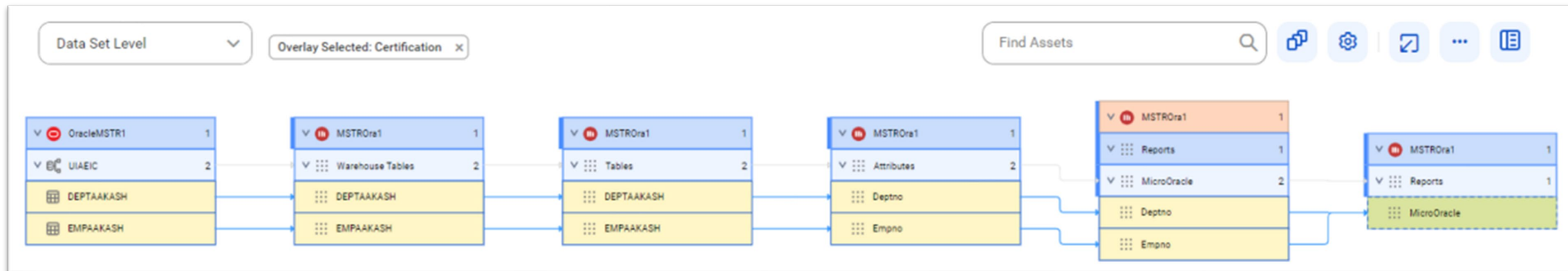
Name ↑	Description
ABA_ROUT	-
ADDRESS	-
BIRTHDATE	-
CC_CVC	-
CC_EXPIREDATE	-
CC_TYPE	-
CCN	-
CITY	-

Scanner Updates (BI Systems)

Business Intelligence and Big Data Analytics

You can harvest metadata describing platform specific assets like reports, documents, looks and corresponding measures or attributes from:

- MicroStrategy
- IBM Cognos
- Google Looker

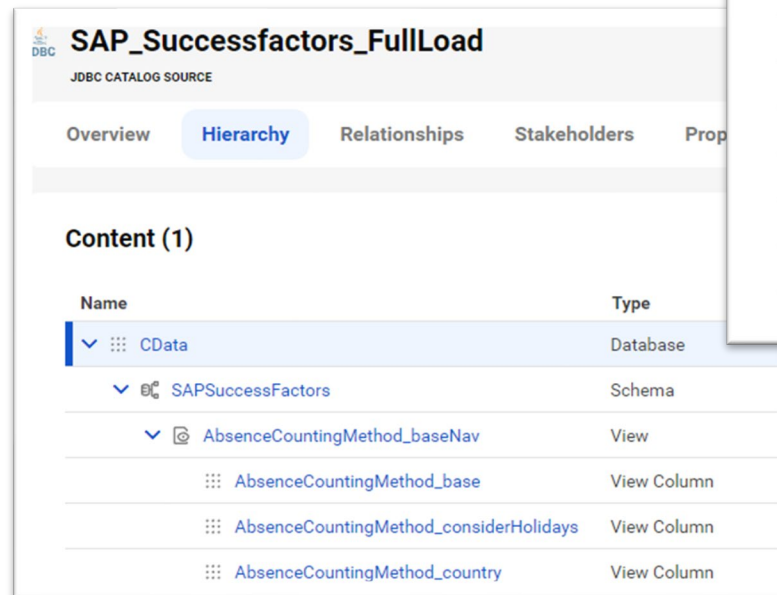


Scanner Updates (JDBC)

JDBC certification

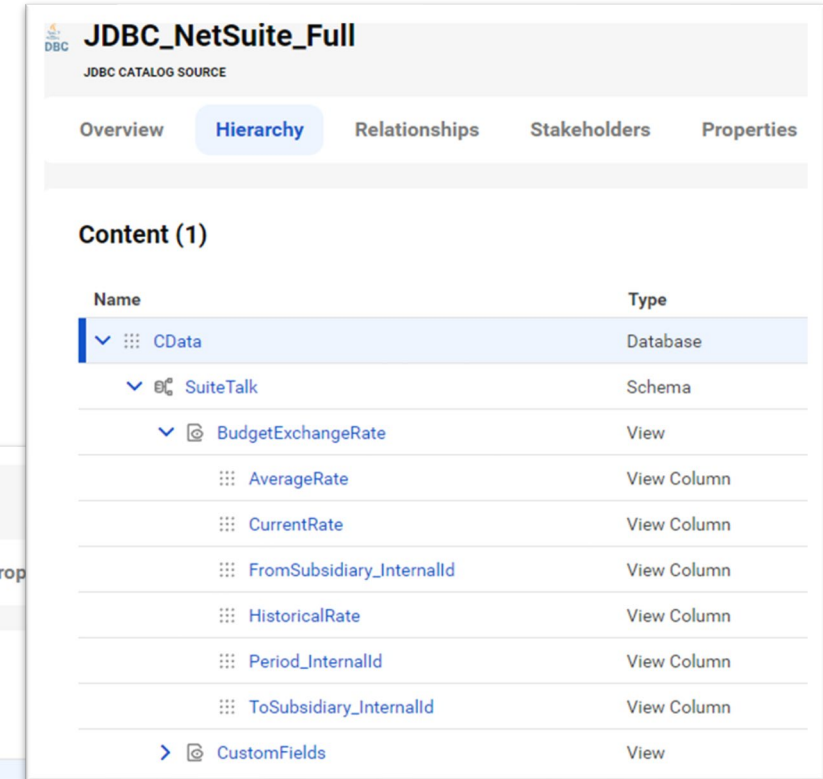
You can use the JDBC catalog source to extract metadata from the following source systems that use certified JDBC drivers:

- Apache Cassandra
- SAP SuccessFactors
- NetSuite



The screenshot shows the 'SAP_Successfactors_FullLoad' JDBC Catalog Source in the Hierarchy view. The 'Content (1)' table lists the following items:

Name	Type
⌵ ⋮ CData	Database
⌵ 📁 SAPSuccessFactors	Schema
⌵ 📁 AbsenceCountingMethod_baseNav	View
⋮ AbsenceCountingMethod_base	View Column
⋮ AbsenceCountingMethod_considerHolidays	View Column
⋮ AbsenceCountingMethod_country	View Column



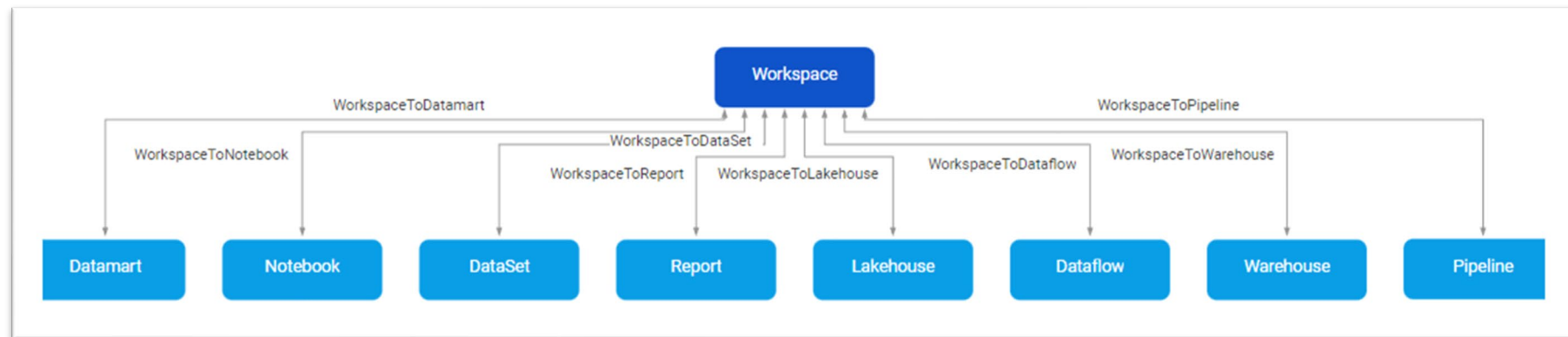
The screenshot shows the 'JDBC_NetSuite_Full' JDBC Catalog Source in the Hierarchy view. The 'Content (1)' table lists the following items:

Name	Type
⌵ ⋮ CData	Database
⌵ 📁 SuiteTalk	Schema
⌵ 📁 BudgetExchangeRate	View
⋮ AverageRate	View Column
⋮ CurrentRate	View Column
⋮ FromSubsidiary_InternalId	View Column
⋮ HistoricalRate	View Column
⋮ Period_InternalId	View Column
⋮ ToSubsidiary_InternalId	View Column
⌵ 📁 CustomFields	View

Scanner Updates (DI based Custom Catalog)

IICS DI based custom catalog source

- You can populate custom model from any source for which IICS has a connector
- Complete conversion of the data from the source to the custom model can be designed/implemented graphically using sync tasks or more complex mapping tasks in CDI



Cloud Data Marketplace Summary of April Release Features

Informatica delivers Cloud-First, Cloud-Native, intelligent, governed Data Sharing for the Enterprise

The screenshot displays the Informatica Cloud Marketplace interface. At the top, the header includes the Informatica logo, "Cloud Marketplace", and a user profile "mmodi-gastaging-fromselfservice". A left-hand navigation menu lists options: New, Home, Search, Data I Own, Tasks, Settings, History, My Orders, Data I Use, Bulk Create, and Dev only features. The main content area is titled "Cloud Marketplace" and contains a welcome message: "Welcome to the Cloud Marketplace! Publish data collections that other teams in your organization can use. Similarly, find and request access to published data collections on a self-service basis." Below this, the "Recent Data Collections" section features four data collection cards, each with a progress bar for requests:

- Customer Profile Data**: Data related to the demographics of a customer, e.g., REQUESTS: 1 Requested
- Customer Contact & Profile Data**: Data related to customers basic information inc. address. REQUESTS: 1 Approved, 1 Fulfilled
- Cust Delivery Preferences**: Customer delivery address, preferred neighbour, secret... REQUESTS: 1 Requested, 1 Approved
- Customer Complaints Records**: Records of past complaints and issues raised by any specific... REQUESTS: 2 Approved

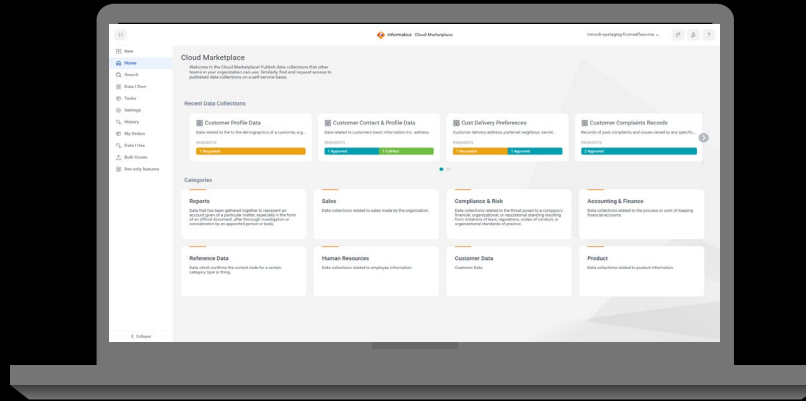
The "Categories" section below lists various data collection types with brief descriptions:

- Reports**: Data that has been gathered together to represent an account given of a particular matter, especially in the form of an official document, after thorough investigation or consideration by an appointed person or body.
- Sales**: Data collections related to sales made by the organization.
- Compliance & Risk**: Data collections related to the threat posed to a company's financial, organizational, or reputational standing resulting from violations of laws, regulations, codes of conduct, or organizational standards of practice.
- Accounting & Finance**: Data collections related to the process or work of keeping financial accounts.
- Reference Data**: Data which confirms the correct code for a certain category, type or thing.
- Human Resources**: Data collections related to employee information.
- Customer Data**: Customer Data
- Product**: Data collections related to product information.

A "Collapse" button is visible at the bottom left of the main content area.

Informatica Cloud Data Marketplace

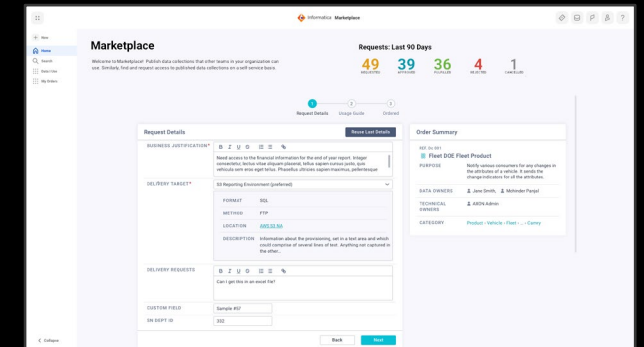
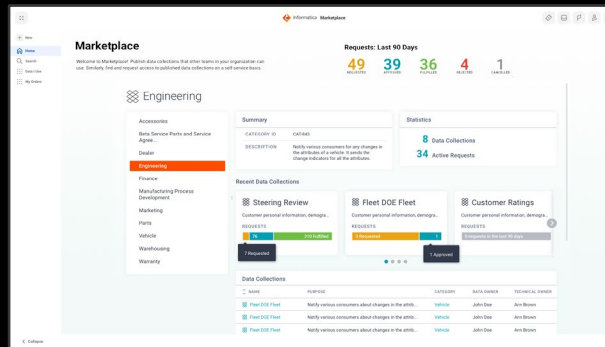
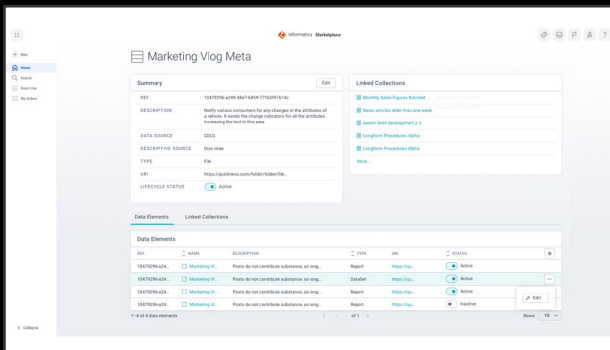
Cloud First/Cloud Native, intelligent, enterprise-scale, governed Data Sharing



Package Diverse Data Assets

Convenient Data Shopping

Automated Delivery & Fulfillment



April Release

- Open Data Quality
 - DQ from any source
 - Import via Public APIs
- Data Collection Variants & Comparison
 - Add relationships between similar Data Collections
 - Compare Data Collections to find the best option for you!
- Star Owners
 - Set thresholds for average approval / delivery speeds of Owners
 - Highlight 'super owners' throughout CDMP
- Other notable features
 - Localization -- German, Japanese
 - Schema Search for Data Assets

The screenshot displays a data quality dashboard with several components:

- Data Quality Rules on Customer Data:** A table with columns: RULE NAME, PRIMARY DATA ELEMENT, DIMENSION, LATEST SCORE, and SCORE DATE. It lists rules like 'Fleet DOE Fleet Pro...', 'ISO Standards', and 'Battery Rollout' with star ratings.
- Super Owners Card:** A card titled 'Super Owners' with the text 'Encourage owners to approve data quickly, star icon in their avatar!' and 'Pending Approval criteria: 0 days'. It also shows 'Fulfillment criteria: 0 days'.
- Add Data Assets to Customer Churn Prediction Pipeline Package:** A modal window with a search bar (set to 'snowflake_retail') and a table of data assets. The table has columns: NAME, DESCRIPTIVE SOURCE, SOURCE PATH, and STATUS. Assets include APAC Sales Data, EMEA Sales Data, LATAM Sales Data, NA Sales Data, Online Sales Data, Returns Data, Store Sales Data, and Subscription Data. The NA Sales Data row is selected. A tooltip shows the source path: 'snowflake_retail_dwh://EMEA_DEMO_ENV/CONTOSO/BKP_DIM_CUSTOMER~com.infa.odin.models.relational.Table'. Below the table, 'Selected Data Assets' shows 'EMEA Sales Data', 'LATAM Sales Data', and 'NA Sales Data' with close buttons. At the bottom, there are 'Cancel' and 'Add 3' buttons.



Demo

Thank you



Cloud Data Quality and Profiling

What's New?

- Exception Management
- Expansion of source types supported by data profiling
- Observability: Scorecard updates
- Verifier updates
- New Data Quality Bundle

Data Quality Exception Management

Use Case

Use Case:

- Ability to identify and view the data that does not meet a data quality criteria

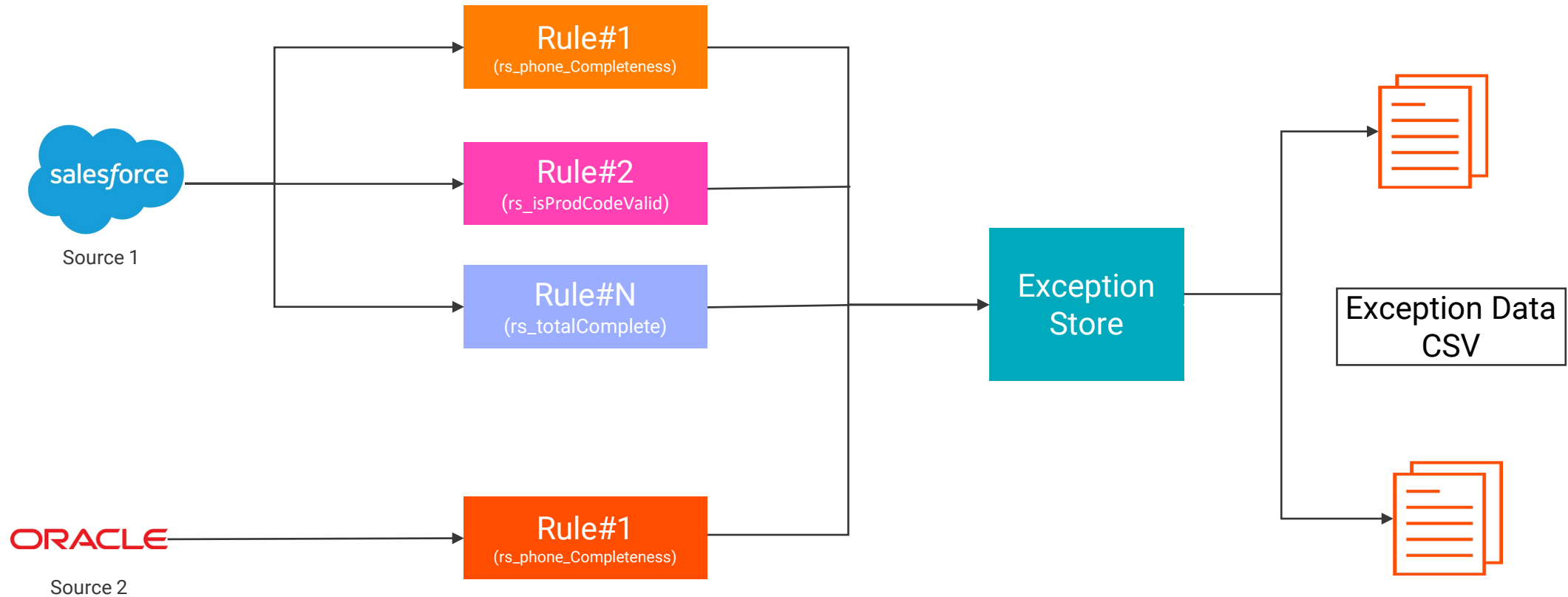
Example:

- Mr. Bob is a DQ Steward working at a multi-national e-Commerce company. He needs to ensure if the quality of the data meets required norms to run the business successfully.
 - For example, users billing information details like name, address, phone, email, credit card, etc., needs to be valid and complete.
- He works with his colleague Ms. Maria, one of the key DQ analyst to ensure the DQ process is implemented **to identify data exceptions** and make the data fit for purpose.

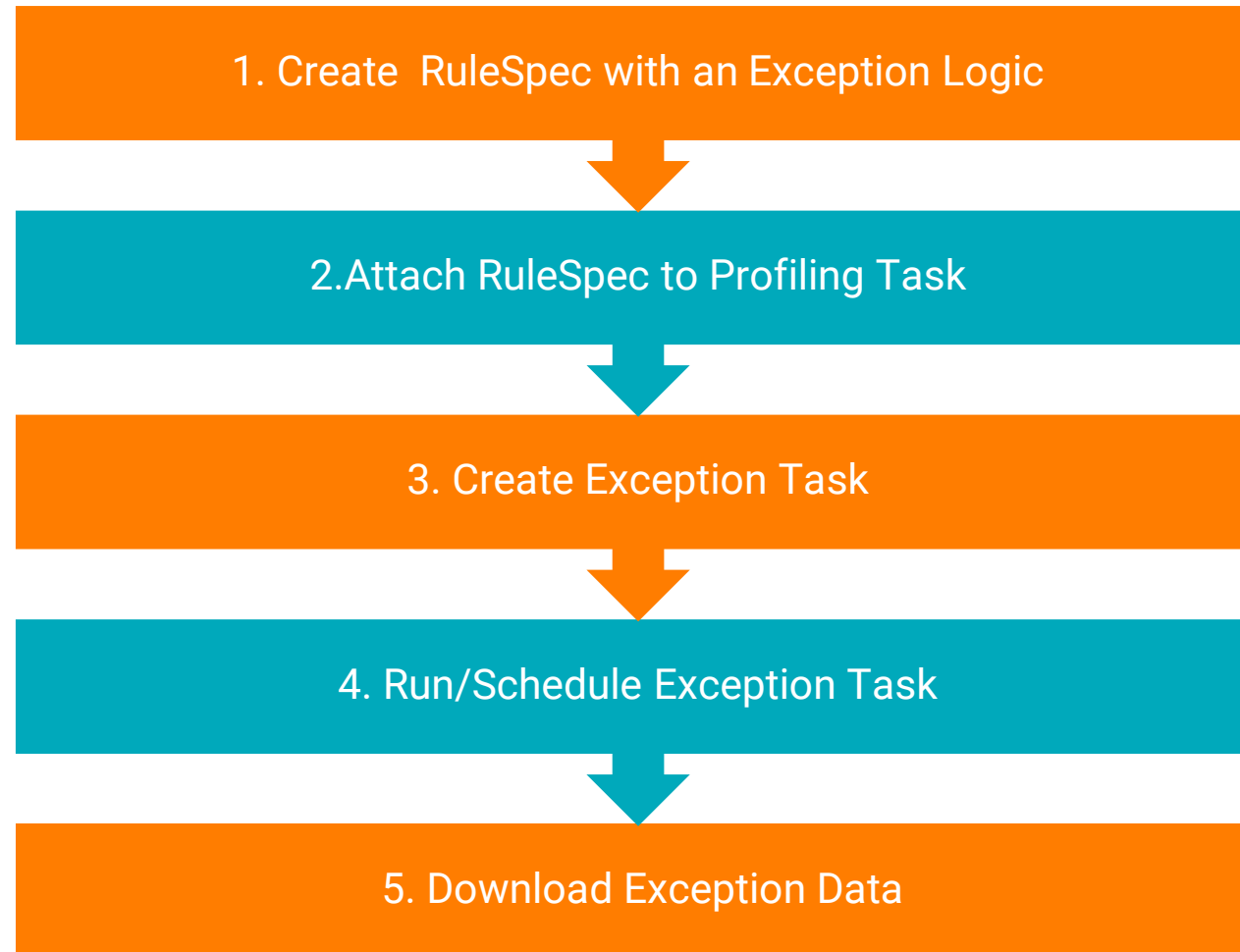


Data Quality Exception Management

Use Cases



Steps to Identify and Download Exceptions Data



Demo

Cloud Data Profiling

What source can you profile?



- 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)
- Snowflake (using native V2 driver)
- Google BigQuery
- Delimited File on: File system | Amazon S3 | Azure ADLS Gen2 | **Google Cloud Storage**
- ODBC (PostgreSQL, Hive and others)
- **SAP ECC, SAP S/4 Hana and SAP BW**
- Databricks Delta tables using ODBC drivers
- Salesforce
 - Salesforce Sales Cloud
 - Salesforce Service Cloud
 - Salesforce Verticals:
 - Salesforce Financial Cloud
 - Salesforce Health Cloud
 - Applications on Force.com
- PARQUET, AVRO and **JSON** files on:
 - Amazon S3
 - Azure ADLS Gen2[Requires CDI-E – profiling runs on Spark]

Deprecated

- Azure Synapse, Snowflake and Redshift using ODBC (still works, not recommended)

Cloud Data Profiling

SAP Connectors

SAP ECC 6.0
All Enhancement
Packages

SAP S/4 HANA

SAP BW / NetWeaver 7.X

Source Details

Connection:

Source Object:

Source Details

Connection:

Source Object:

Source Details

Connection:

Source Object:

Demo

Observability: Data Profiling

Alerting/Notifications for scorecards

Use case:

- Notify user when a Data Quality score worsens
- Allows monitoring of scores without checking scorecards in UI

Features:

- Triggered when scores changes from:
 - Good -> Acceptable
 - Acceptable -> Not Acceptable
 - Good -> Not Acceptable

© Informatica. Proprietary and Confidential.

The screenshot shows an email notification titled "Data Quality Failure Notification for NotificationRO". The sender is "noreply@informaticacloud.com" and the recipient is "Kumar, Nilesh". The email includes the Informatica logo, the date "2023-01-30 19:05:33 IST", and a retention policy of 3 years. The main message is "Data Quality Status change from Acceptable to Not Acceptable". Below this, it lists the rule occurrence name as "NotificationRO", the description as "NotificationRO", and the rule dimension as "Accuracy". A legend for "Rule Occurrence Thresholds" shows three categories: Not Acceptable (red, 0-60%), Acceptable (yellow, 60-80%), and Good (green, 80-100%). Two scorecards are shown: the "Current Score" is 20.0 (Not Acceptable) with 100 total rows and 10 failed rows, and the "Previous Score" is 60.0 (Acceptable) with 94 total rows and 9 failed rows. The score change is a decrease of 66.67. At the bottom, it identifies the "Primary Data Element" as "NAME" and provides a link to "Open Data Quality Rule".

Observability: Data Profiling

Alerting/Notifications for scorecards

Features:

- **Notification recipients:** all stakeholders of metrics/Rule Occurrences
- In app and email notification
- Ability to:
 - Turn notifications on/off at stakeholder/IDMC user level
 - Receive Realtime notifications or digest

The screenshot displays the Informatica Data Governance and Catalog interface. The main window shows the 'Settings' page for 'Data Governance and Catalog'. A table titled 'Notification Settings' allows users to configure notifications for various services. The 'Data Quality' service is highlighted, showing that both 'Email Summary' and 'Email Event' notifications are enabled. To the right, the 'Notification Summary Email Settings' section shows the 'Email Summary Frequency' set to '24 Hours'. An inset window titled 'Notifications' shows a list of recent alerts, including status changes from 'Good' to 'Not Acceptable' for various data quality rules, with timestamps and links to 'Open Rule'.

Service	Notification Type	Email Summary	Email Event
Collaboration		<input type="checkbox"/>	<input checked="" type="checkbox"/>
Data Quality		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
Ticket Manager		<input type="checkbox"/>	<input checked="" type="checkbox"/>
User Jobs		<input type="checkbox"/>	<input checked="" type="checkbox"/>
Workflow Tasks		<input type="checkbox"/>	<input checked="" type="checkbox"/>

Verifier Upgrade

- All new Address Verification features will be available with latest version:
 - Expanded writing script support (ie.: Arabic support)
 - Reverse Geocoding
 - QuickCapture (purpose-built type-ahead, address look-up functionality)
- Additional features and outputs can now be added.
 - Granular data quality indicators like Element Result status on the roadmap
 - Verification Status Code with one character, V, A, C, I
 - Address **v**alid, was data **a**dded, was data **c**orrected, is address **i**nvalid.
 - Result Quality using a scale of 0 to 6
 - Describes the quality of the address and replaces the numeric digit contained in the former status code.
 - 6 is highest quality
 - 0 is lowest quality
 - Residue ports

© Informatica. Proprietary and Confidential.

Verifier Upgrade

Input and output ports changes

- **Result Quality** and **Address Type** two new ports added inside Status Codes output.
- **Country ISO Number 1** and **Country ISO Number 2** are two new ports added inside Country output port.
- Enrichments > Geo Coding updated with new sections, **Arrival Point**, **RoofTop**, **Street Center**, **Locality Center**, **Postal Code Center**.
- **Certification Status** and **Country Specific Status** are now moved out from the country level status to individual common status.

The screenshot displays the Verifier configuration interface with several sections highlighted by red boxes:

- Enrichments** (expanded):
 - Geo Coding
 - Arrival Point
 - Arrival Point Accuracy Code
 - Arrival Point Coordinates
 - Arrival Point Latitude
 - Arrival Point Longitude
 - RoofTop
 - Rooftop Accuracy Code
 - Rooftop Coordinates
 - Rooftop Latitude
 - Rooftop Longitude
 - Street Center
 - Street Center Accuracy Code
 - Street Center Coordinates
 - Street Center Latitude
 - Street Center Longitude
 - Locality Center
 - Locality Center Accuracy Code
 - Locality Center Coordinates
 - Locality Center Latitude
 - Locality Center Longitude
 - Postal Code Center
 - Postal Code Center Accuracy Code
 - Postal Code Center Coordinates
 - Postal Code Center Latitude
 - Postal Code Center Longitude
- Country
 - Country Name 1
 - Country Name 2
 - Country ISO3 1
 - Country ISO3 2
 - Country ISO2 1
 - Country ISO2 2
 - Country ISO Number 1
 - Country ISO Number 2
- Certified
 - Certification Status
 - Australia-AMAS
 - Canada-SERP
 - France-SNA
 - New Zealand-Send Right
 - United States-CASS


Input Country Port

- Country
 - Country

- Status Codes
- Verification Status Code
- Result Quality
- Match Percentage
- Address Type
- Country Specific
- Country Specific Status
- Australia
- Austria
- Belgium
- Brazil
- Czech Republic
- France
- Germany
- Italy
- Japan
- Poland
- Russia
- Serbia
- South Africa
- Spain
- Switzerland
- United Kingdom
- United States

Data Quality Financial bundle

- Pre-configure Data Quality assets focused on Financials
 - Bank related codes
 - Transit Number, credit card codes, Account number validations
 - Rating codes
 - Standard & Poor's, Moody's
 - Stocks
 - CUSIP validation

 Data Quality Financial Bundle Copy Bundle Content To...

View available bundle

Bundle Details

Name:	Data Quality Financial Bundle
Description:	Use the Financial Bundle to verify and enhance data from organizations in the financial services sector. The Financial Bundle includes rules that perform the following data quality processes, contact data cleansing, financial data cleansing, general data cleansing, matching and deduplication. The bundle includes some rules from Core, GBR and Americas Bundles.
Publisher:	DQ Content Team
Version:	1.0.1
Bundle Type:	Public
Allow:	Copy
Published On:	Feb 7, 2023, 12:06:13 AM

Demo