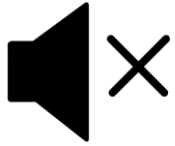Nov 28, 2023

What's New in Informatica Cloud Data Governance and Catalog (CDGC), Cloud Data Marketplace (CDMP) and Cloud Data Quality (CDQ)? - October and November 2023 Release

- Amrish Thakkar, Sr Director, Product Management

- Ricardo Sobral, Sr Director, Product Management

- Kate Amory, Principal Product Manager, 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 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.
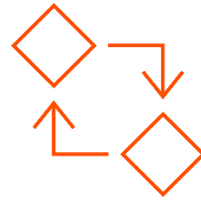
Informatica

Feature Rich Success Portal
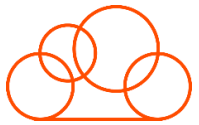
Bootstrap trial and POC Customers

Enriched Customer Onboarding experience
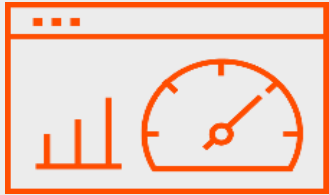
Product Learning Paths and Weekly Expert Sessions
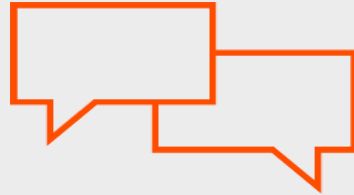
Informatica Concierge

Tailored training and content recommendations
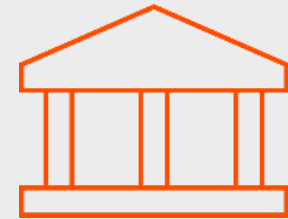
More Information

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 2023.11.S

- **Auto Cataloging of Informatica Cloud Data Integration (CDI) (Preview)**

- **Informatica QuickLook for Google Chrome**

- **Custom Workflow Support including Ticket Management Experience**

- **New Asset Types**
  - Legal Entity, Geography, Business Area, Project, Regulation

- **Asset Page Customization**

- **Data Classification enhancements**
  - Deleting Data Element Classifications
  - Labelling Low Sensitivity
  - Glossary Association with Data Element and Entity Classifications
  - New OOTB Classifications

- **Granular Stakeholder assignment for Technical Assets**

- **Usability Improvements**
  - Lineage
  - Bulk Import (Tool Tip and Instruction Sheet)
  - Show list of Users in a User Group
  - Cloning Catalog Source

- **Support for Multiport DQ rules (Manual Rule Creation)**

- **Scanner Additions and enhancements**

- **Profiling**

- **Discovery Updates**

Auto Cataloging of IDMC Metadata (Preview)

Auto Cataloging of IDMC Metadata (CDI Only)

Feature description

- Key highlights

  - Automatically catalog CDI design time metadata as changes are made to mappings, tasks, etc.

  - Automatically derive lineage from CDI tasks as they get executed.

  - Support for task execution part of TaskFlows

  - Monitor/retry processing of runtime metadata in MCC for each individual tasks

- Benefits

  - Easy configuration for cataloging asset part of IDMC

  - Easier adoption of CDGC for existing CDI customers

  - Foundation infrastructure for automated cataloging of all IDMC Services

Auto Cataloging of IDMC Metadata config

- Simply enable in MCC

- Select the runtime environment to use for runtime metadata processing

- Allow defining filter to include which Project/folders to scan

- Allow define expert parameters if required

Proj_taskflow

PROJECT

LIFECYCLE
PUBLISHED

LAST UPDATED
Nov 6, 2023, 4:06 AM

Overview | Hierarchy | **Relationships** | Stakeholders | Attributes | Tickets | History

Find

- 1 — IDMC Resouce
- 2 — Mapping
- 1 — Linear Task Flow
- 1 — PowerCenter ta...
- 1 — Synchronizatio...
- 11 — TaskFlow
- Proj_taskflow
- 3 — Mapping task
- 3 — Informatica Dat...
- MT_SourceTarg...
- MT_SourceTarg...
- MT_LookupTable
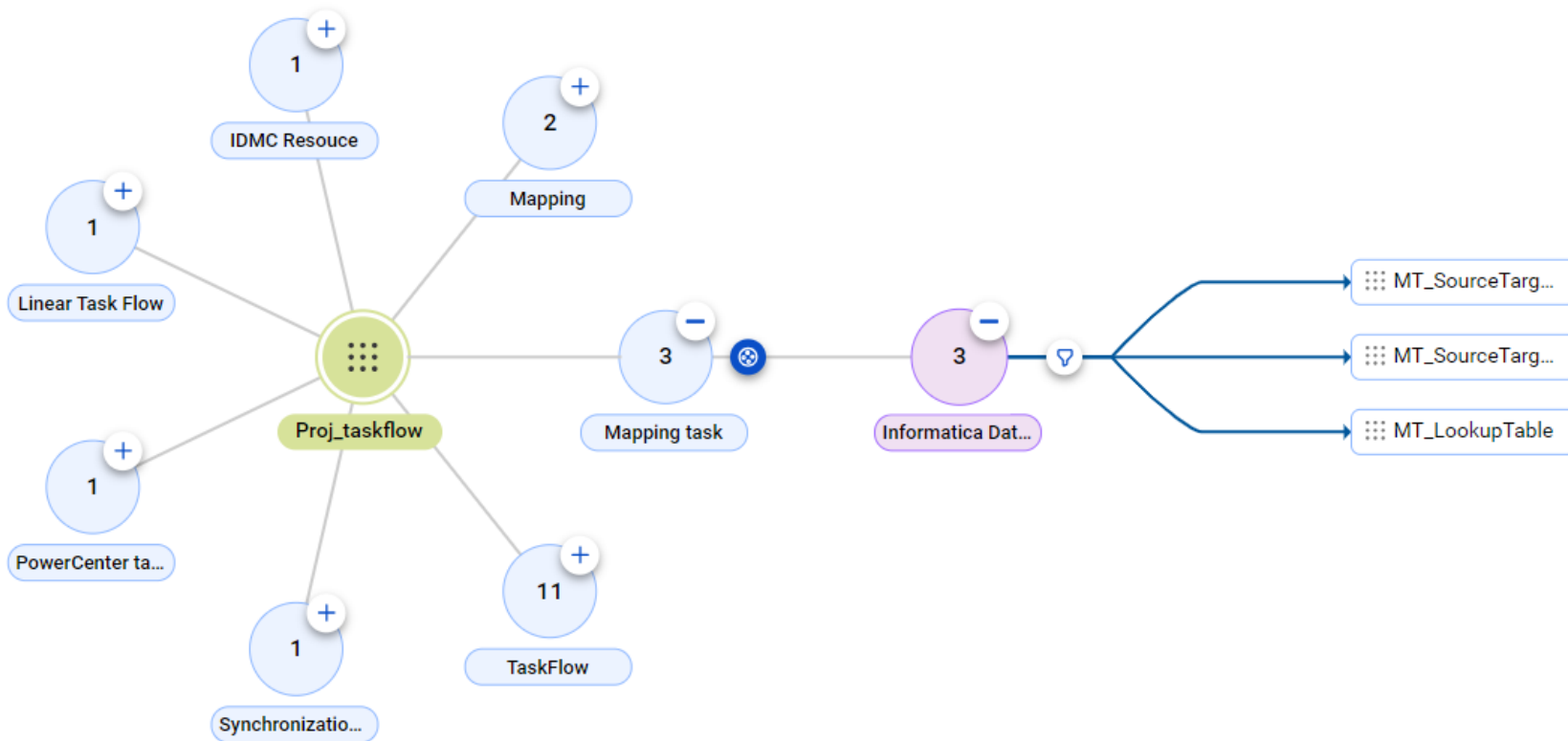
IDMC Catalog Source vs Auto Cataloging

Comparison as of 2023.11

IDMC Catalog source	Auto Cataloging
CDI (excluding Adv.), B360 Metadata assets - Mappings, tasks, Business entities	CDI (excluding Adv.) - Mappings, tasks, taskflows
Metadata from any Org	Only CDGC Org Metadata
Historical runtime metadata (up to 18 months)	Only lineage for new tasks execution
Synchronization in batch (on schedule)	Synchronization in (near) realtime
Tasks executed via Taskflow are with workaround	Tasks executed via Taskflow are supported
Parameter file are supported	Parameter files are **not** supported
Connection aware scan is supported	Connection aware scan supported
Custom attribute **can** be associated with IDMC Assets	Custom attribute **cannot** be associated with IDMC Assets
Stakeholder can be defined granularly	Stakeholder can be only defined at top level

Informatica®

Asset Page Customization

Asset Page Customization

Customise Assets Views for different constituents

Key Highlights

- Create different views of objects to better provide the information different users need

- Add/remove/rename fields / object tabs

- Change the appearance / order of fields

Benefits

Increased engagement:

- Users find relevant information more easily

- Re-ordering promotes the things they value the most

Informatica

Feature Overview

- Ability to design a custom view of an object for each asset type and assign the view to specific roles or user groups.

    - Improves the usability as this allows for Asset views based on a Persona's requirements

- Depending on the roles and user groups, user may have access to multiple custom views

    - Users can select a default view and switch between views in the UI.

- All users still have access to the OOTB layout for all object types

- Configuration: An additional feature privilege is available in MCC to create custom views of asset pages.

Informatica

Different Custom Views of the Same Object

Views above use different column layout for Overview Tab.

View to the right has fewer tabs, reordered layout with custom panels and naming

Informatica QuickLook: Browser Extension

Informatica QuickLook

Browser Extension for Google Chrome

- QuickLook provides an Ability to quickly search for data assets available in the catalog without having to open the Cloud Data Governance and Catalog application.

- Ability to seamlessly search for text on any web page to find the corresponding assets in your catalog.

- Enhanced browsing experience by conveniently viewing or previewing content without navigating away from the current web page.

- 3 options available to search for the items:

    - Keyboard Shortcut and Right Click options are available after Text selection

    - User can search any item by opening QuickLook

- Available only for **Google Chrome** currently and can be installed from Chrome store

Informatica®

Screenshots for QuickLook

Custom Workflow & Ticket enhancements

Custom Workflow

Feature description

- Key highlights

  - Create complex Multi Step Approval workflow based on your organizational needs

  - Support for Human Task, Exclusive / Inclusive Gateway

  - Assign Custom Workflows based on asset types and roles

  - Collaboration on Tickets (Discussion , Comments)

- Benefits

  - Complex workflows can be created in a Canvas using Drag and drop

  - Extend Governance customization to support broader deployments

  - Custom workflows in line for organization policies

Informatica
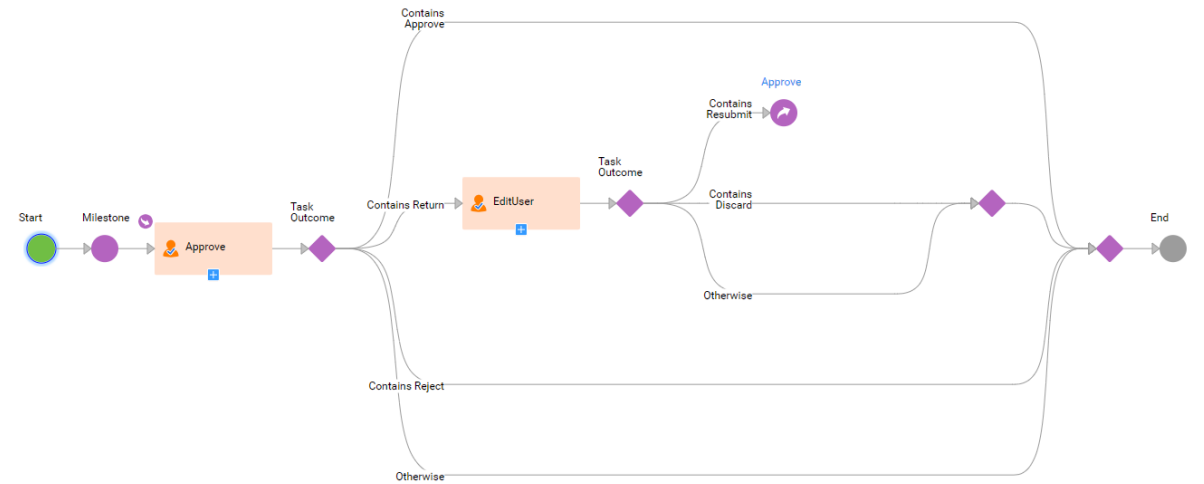
Workflow functionality supported

- Streamlined Approval Process – Single / Multi Step Workflows with Human Task, Exclusive Gateways

- Granular control on task by assigning relevant stakeholder role to act on each task

- Real Time Workflow Progress update on tickets on Current , Completed and Not Completed Tasks

- Audit trail for all approval activity on the ticket

Informatica

Screenshots for Workflow

Single Step Approval

Ticket - Task Status

Task Assignment

Multi Step Approval

Informatica

List of OOTB Workflows available

- Out-of-Box Workflow available as "Data Governance and Catalog Workflow Bundle" in Administrator

- Following workflow are available as OOTB

  - ManageAsset_SingleStep Approval Workflow

  - ManageAsset_TwoStep Approval Workflow

  - ManageAsset_ThreeStep Approval Workflow

New Asset Types

New Business Asset Types

- Five new Business Types has been added as part of this release:

| Legal Entity | Geography | Business Area | Project | Regulation |

- All these assets can be represented hierarchically

- You can add Custom Attributes as needed to extend the model

- These assets can be created/updated from UI or using Bulk Import

- Role configuration has been updated to allow for Create, Read, Update, Delete privilege to be granted via the role

  - OOTB Roles have been updated to grant appropriate privilege

Informatica

Relationships

- All new assets are hierarchical to help organize the content

- Main goal is to assist search via creation of relationships

- Relationships Matrix will be published with release

- Legal Entity / Business Area / Geography are contextual assets (centrally managed lists, relationships made to them by the community)

    o Permitted relationships are meant to help contextualize main data oriented assets

    o Note that Geography is more connectable, reflecting e.g. privacy use cases – locations of systems/ data sets

    o 'Direction' of relationship to make it easier for owners to add context

Informatica®

Other Enhancements

Granular Stakeholder Assignment for Technical Assets

- Until now, you could assign Stakeholders only at the Catalog source and all child assets will inherit from the Catalog Source.

- With this release, Stakeholders can be assigned at any level in the hierarchy for Catalog Sources

  - Making any change to a stakeholder for an Asset will stop inheriting any stakeholders from the parent

  - Asset will always inherit from the immediate parent

  - UI indicates whether the stakeholders are inherited or Manually assigned

- Ability to assign technical stakeholders at different levels of hierarchy from UI or Bulk Export/Import.

Catalog Source
Manually Assigned
Data Steward: John
Data Owner: Mary

Databased
Inherited from Catalog Source
Data Steward: John
Data Owner: Mary

Schema
Manually Assigned
Data Steward: Craig
Data Owner: Bill

Table
Inherited from Schema
Data Steward: Craig
Data Owner: Bill

Column
Inherited from Table
Data Steward: Craig
Data Owner: Bill
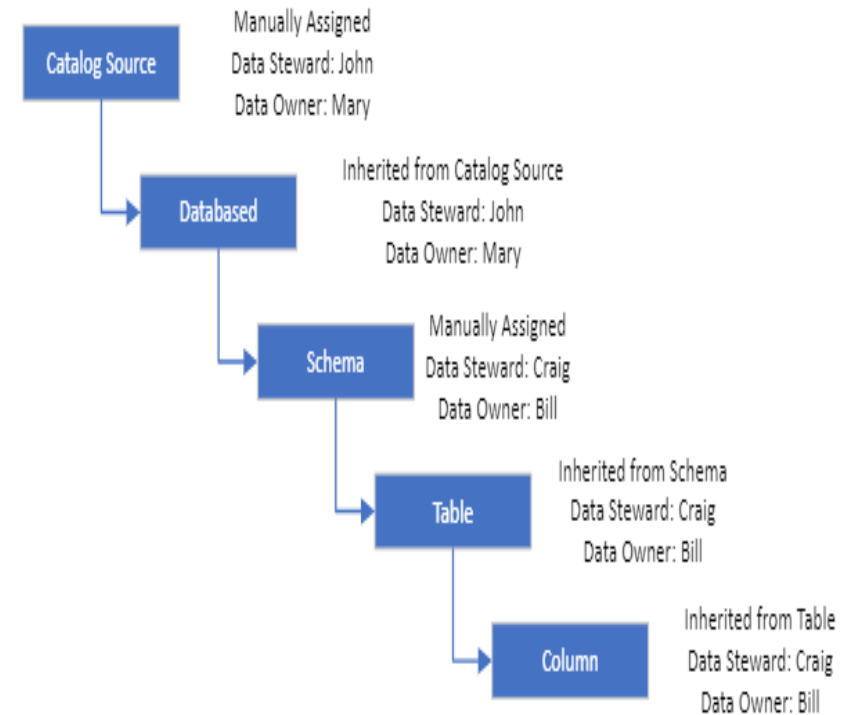
Inherited v/s Specific Assignment

Inherited Assignment of Stakeholders

- By default the Stakeholders will be inherited from the immediate parent and will be marked as such

Specific Assignment of Stakeholders

- Modifying any stakeholders that have been inherited are shown as "Specifically Assigned"

Display list of Group Users

- With this release, users can now see the list of users in the Group if the Group has been assigned as a Stakeholder

- Users can explicitly expand the group to see the list of users

Lineage Enhancement

We are adding following options to Lineage :

- Expand All

  - Full lineage that will load progressively

    - Progressive loading optimizes the user experience, ensuring smooth navigation through the lineage diagram even in scenarios with large and complex lineage.

  - User will have option to stop loading at any point

  - Users can explore lineage diagrams more efficiently with the 'Expand All' option, saving time and effort in navigating.

- Show legend option in lineage diagram saves user time by providing quick access to legend whenever required.
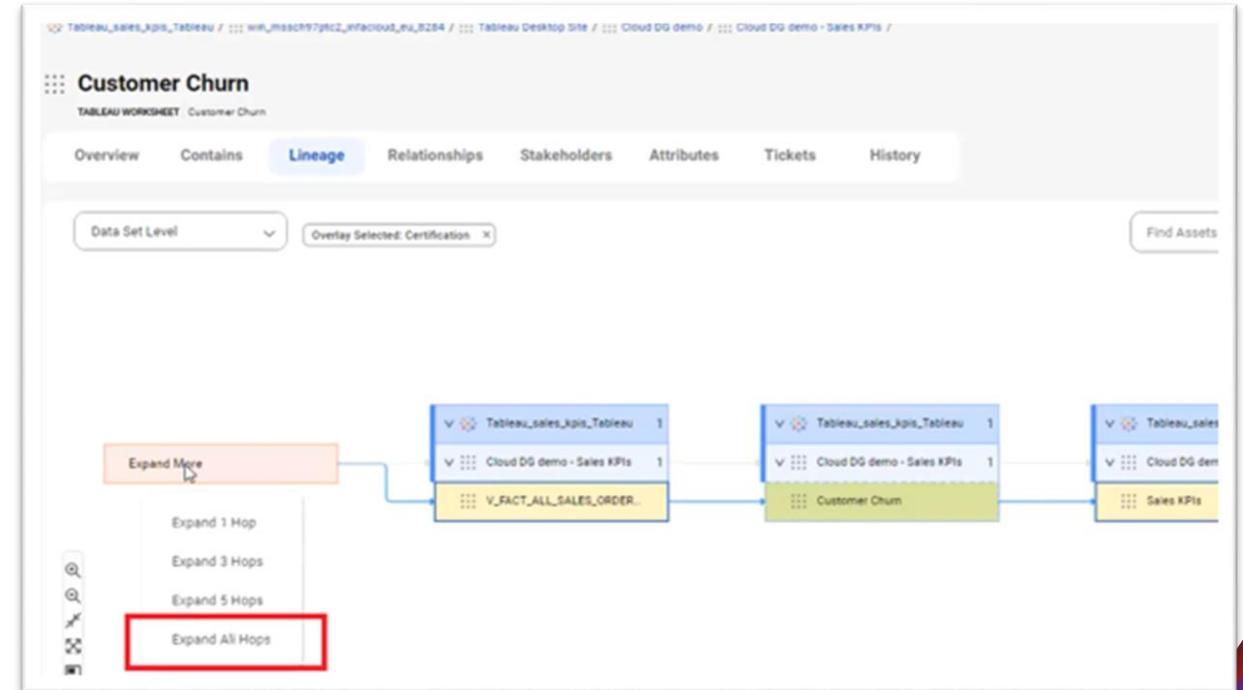
Classification Enhancements : Sensitivity - None

- Data Element Classification Sensitivity level with "None" option

    - "None" sensitivity level provides a standardized approach for classifying data without specific sensitivity requirements.

    - Seeing low classification sensitivity directly in the CDGC UI and Lineage diagram aids in decision-making regarding data sensitivity.

Classification Enhancements : Glossary Relationship

- Link Data Element Classification & Data Entity Classification with Business Glossary (Business Terms and Metrics) using relationship.

  - Establish a relationship between Data Element and Data Entity classifications with the Business Glossary (Terms and Metrics).

  - After Execution of the Catalog Source Glossary Association Job, glossaries are automatically linked to data elements and data sets where classifications are associated. This automation simplifies the curation process, ensuring accurate and efficient associations.
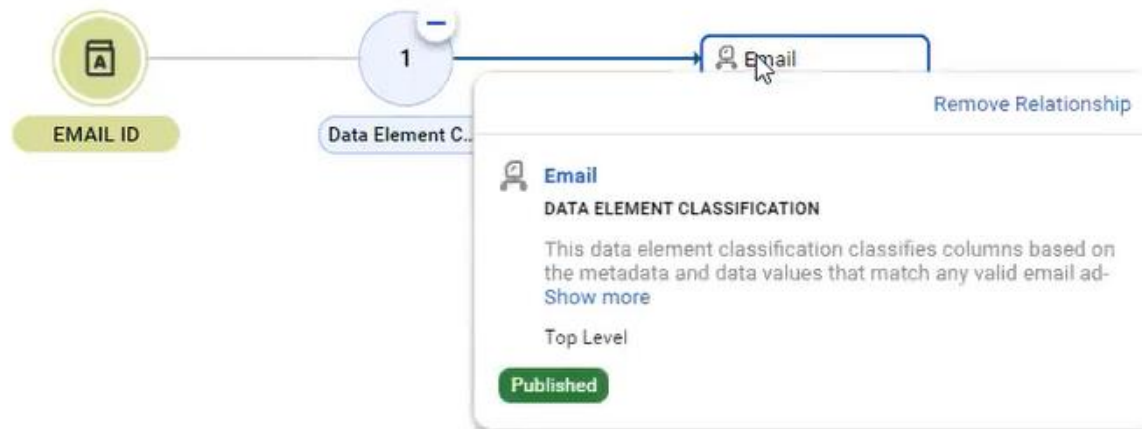
Clone Catalog Source Configuration

In this release we are allowing to clone a catalog source config in MCC.

- The cloning feature enhances operational efficiency by allowing users to replicate configurations with minimal possible efforts. This is particularly useful in scenarios where similar configurations are needed across multiple catalog sources.

  - In some cases while dealing with large catalog sources, the ability to split catalog sources becomes imperative. This enhancement facilitates administrators in efficiently replicating configurations and just change filters as needed.

- Only configuration is cloned. User has the option to rename the cloned Catalog Source before saving
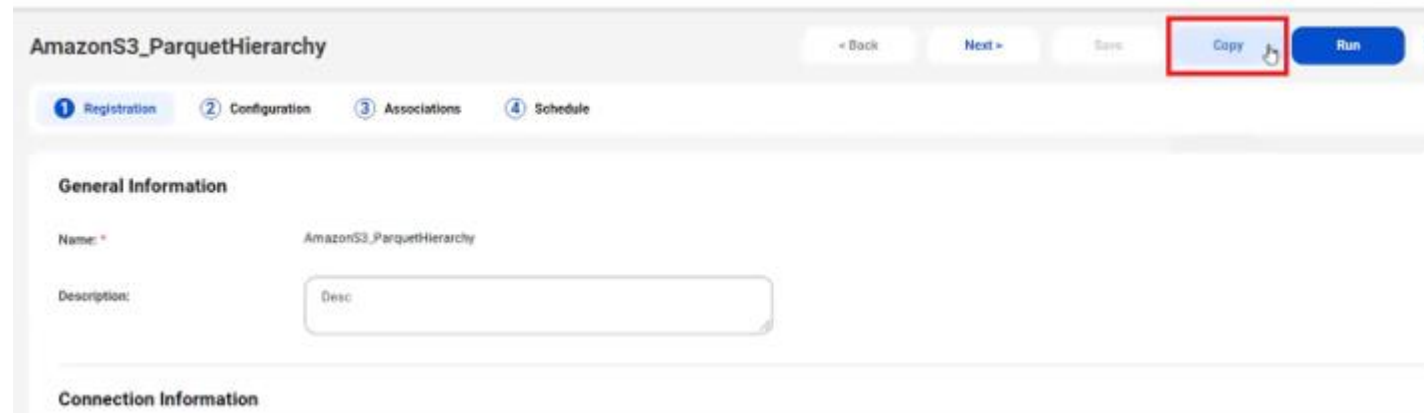
Multi Input Port DQ Rule

- Support Multi Input Port DQ Rule has been added

  - It enables evaluating Validity for a field based on inputs from multiple fields

- Ability to select and Execute Multi Input Port CDQ Rule in Rule Occurrence

- User need to map Primary and Secondary data element to CDQ Input Ports in Rule Occurrence – Port mapping is mandatory for multi-port rule

  - Both Primary and Secondary Data element for a Rule Occurrence should be from same table

- Scores will always be reported on Primary Data Element

  - Score Aggregation and Propagation will consider only Primary Data Element

- Multi Input Port Rule is currently supported for manual Rule Occurrence (Not for Rule Template and automation)

DQ_DQRO_MP_CDGC_52228_1698857088943
DATA QUALITY RULE OCCURRENCE

| Overview | Score | Stakeholders | Tickets | History |

Target:	Threshold:
90	70

Technical Rule Reference:	Technical Description:
Default / DQ_MULTIPORT_RULE_CDGC_52228_16988570 88943	DQ_MULTIPORT_RULE_CDGC_52228

Input Parameter	Data Elements
NewInput1	EMPLOYEE_ID
NewInput2	FIRST_NAME

Informatica

Additional Public API

- Following CDGC Public API is added in the release

    - **Catalog Source Management API**

        - **Catalog Source Execution API**

- Using this API, you can execute a catalog source with selected capabilities

    - DELETE / RETAIN option is not changeable in Execution API.

    - Only Saved option / capabilities at Catalog Source will be considered for execution.

    - Snowflake Writeback capability will not be supported.

Profiling Updates

- **Databricks**

  - Custom query support

    - Percentage: *TABLESAMPLE (30 PERCENT)*
    - Fixed number of rows: *TABLESAMPLE (5 ROWS) or LIMIT 2 or LIMIT 2 OFFSET 3*
    - Repeateable sample: *TABLESAMPLE (30 PERCENT) REPEATABLE (123)*
    - Numebr of rows defined by formula: *LIMIT length('SPARK'); (limit with an expression* `to define the limit`)

  - Profiling of delta tables in Unity Catalog

- **PAM Additions**

  - Delimited files in Local and Remote file systems

  - Google BigQuery views

Databricks_TestSampling_NewCS

Enable Data Profiling: 🔵

Connection and Runtime

Runtime Environment: scharishvm.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

Sampling Type: * ❓   Custom Query ⌄

Sampling Query: * ❓   LIMIT 2

Maximum Precision of String Fields: * ❓   50

Text Qualifier: * ❓   Double Quote ⌄

Enter custom query here

Discovery PAM Updates

- Support for Intelligent Glossary Association, Data Classification, Relationship Discovery has been added for the following sources

  - ADLS Gen2

  - Teradata

  - SAP (BW/4HANA, HANA, BO)

  - Azure Data Factory

  - PostGreSQL

- Support for Intelligent Glossary Association, Data Classification has been added for the following sources

  - Tableau

Scanners

New Scanner - Oracle Business Intelligence scanner

Oracle Business Intelligence
A business intelligence tool to gather, store, and visualize enterprise data.
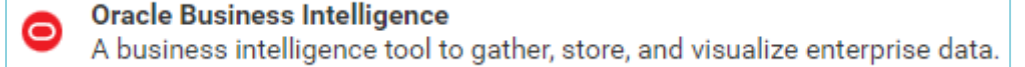
- Oracle Business Intelligence is a new scanner to extract metadata from Oracle Business Intelligence Enterprise Edition Server

- The scanner can work with both Linux and Windows secure agents

Connection Information

Catalog Source Type:	Oracle Business Intelligence
Version: *	OBIEE 11.x
Server URL: *	http://inw12edcs002:9502/analytics/saw.dll
Login User: *	▮▮▮▮▮▮
Login Password: *	••••••••••••••••••••••••••••••••
Metadata File: *	C:\OBIEESamples\Sample17.xml

Configuration Parameters

Variable Values File:	
Data Sources File:	C:\OBIEESamples\DatasoureFile\datasources.xml
Optimize For Large Models:	○ True ● False
Incremental Import:	○ True ● False
Worker Threads:	
Miscellaneous Options:	-m 4G

New Scanner - TIBCO Spotfire scanner

TIBCO Spotfire
A business intelligence tool that you can use to explore data
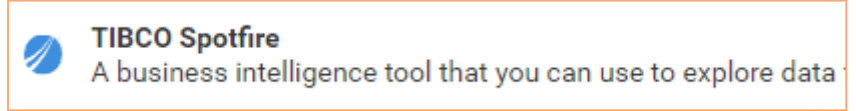
- TIBCO Spotfire is a new scanner to extract metadata from Tibco Spotfire Server

- The scanner can work with both Linux and Windows secure agents

Connection Information

Catalog Source Type:	TIBCO Spotfire
Server URL: * ❓	http://localhost:8080/spotfire
Client ID: * ❓	1a2b3c4d5e6f7g8h91a2b3c4d5e6f7g8|oauth-clients.spotfire.tibco.com
Client Secret: * ❓	••
Download User: ❓	██████████
Download Password: ❓	•••••••••••••••••••••••••••

New Scanner - Qlik Sense Cloud scanner

Qlik **Qlik Sense Cloud**
A cloud-based business intelligence tool to connect to data and create reports.

- Qlik Sense Cloud is a new scanner to extract metadata from cloud-based Qlik Sense spaces, applications and QVDs

- The scanner can work with both Linux and Windows secure agents

Connection Information

Catalog Source Type:	Qlik Sense Cloud
Qlik Sense Cloud URL: * ❓	https://1a2b3c4d5e6f7g8.us.qlikcloud.com/
Token: * ❓	••

Informatica

New Scanner - Oracle Cloud Infrastructure Golden Gate

Oracle Cloud Infrastructure GoldenGate (Preview)
An Oracle data-replication tool to load, distribute, and filter data

- OCI Golden Gage is a new scanner whose task is to scan the metadata of GoldenGage deployments instances and provide lineage between replicated objects.

- In the tech preview version, the scanner supports replication between Oracle ADB and Oracle MySQL databases.

Connection Information

Catalog Source Type:	Oracle Cloud Infrastructure GoldenGate (Preview)
Oracle GoldenGate URL: *	https://████████deployment.goldengate.us-phoenix-1.oci.oraclecloud.c
Oracle GoldenGate User Name: *	██████
Oracle GoldenGate Password: *	●●●

Informatica

Scanner Update - Oracle ADB

- Support for external tables

- Support for dbms_cloud functions: COPY_DATA and EXPORT_DATA.
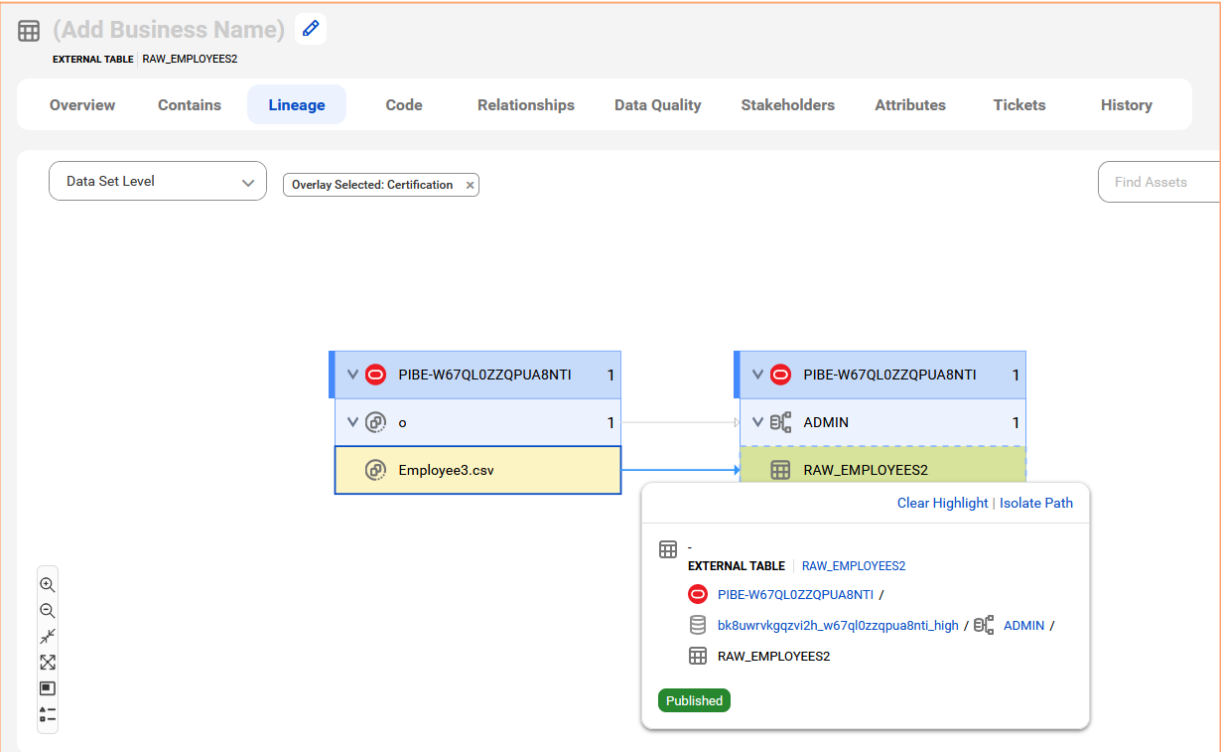
Scanner Update - Microsoft Power BI

- Added Microsoft Power BI connection type in IDMC Administrator

- Support for user defined functions in Power Query scripts

- Support for scanning Paginated Reports

- Support for Dataflows as Power BI reports sources

Microsoft Power BI
A business intelligence tool to connect to data and create reports

Data Set	1
Dataflow	**2**
Dataflow Table	2
Field	16
Report	1
Workspace	1

Connection Details

Connection Name:*   PowerBI_newConnSPN

Description:

Type:* ❓   Microsoft Power BI

Microsoft Power BI Properties

Runtime Environment:* ❓   invrh79cdmp04

Connection Section

PowerBI Cloud Connection Mode:* ❓   Service Principal

PowerBI Cloud URL:* ❓   https://api.powerbi.com/

PowerBI Cloud Auth URL: ❓   https://login.microsoftonline.com

PowerBI Cloud Scope: ❓   https://analysis.windows.net/powerbi/api/.defau

PowerBI Cloud Client ID:* ❓   ▮▮▮▮▮▮▮▮▮▮

PowerBI Cloud Tenant ID:* ❓   ▮▮▮▮▮▮▮▮▮▮

PowerBI Cloud Client Secret:* ❓   ••••••••••

Informatica

Scanner Update - Google Big Query

- Support for native SQL programable objects:

    - user functions

    - stored procedures (GoogleSQL)

Google BigQuery
A serverless data warehouse that enables scalable analysis

```
CREATE PROCEDURE ecosystemqainfa62360.ADVSCANNER_QA.CROSSDATASET_PROCEDURE()
BEGIN
DECLARE sql_stmt STRING DEFAULT "insert into ecosystemqainfa62360.ADVSCANNER_HAWK.MAPPER select STATE, TEMPERATURE, HEADERS
from ecosystemqainfa62360.ADVSCANNER_QA.MATTER FULL OUTER JOIN ecosystemqainfa62360.ADVSCANNER_QA.COMPANY ON
MATTER.TEMPERATURE = COMPANY.EMPLOYEES";
EXECUTE IMMEDIATE sql_stmt;
END
```
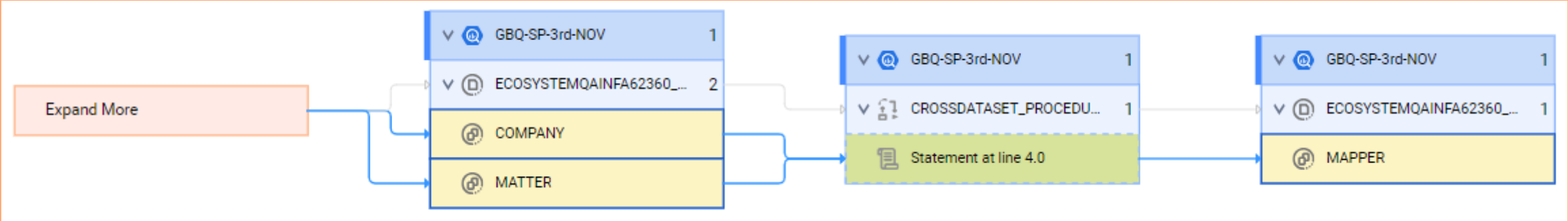
ScannersUpdate – IICS Scanners

- Support for additional connectors

  - Mongo DB Reader/Writer

  - Apache Hive Reader/Writer

  - Business 360 FEP Writer

  - SAP ODP Extractor

  - Sharepoint Reader

Informatica Intelligent Cloud Services
An on-demand subscription cloud service that provides access to Data Integration and Business 360 assets.

Informatica PowerCenter
An ETL tool that is used to extract data from multiple sources, transform, and load them into different target databases.
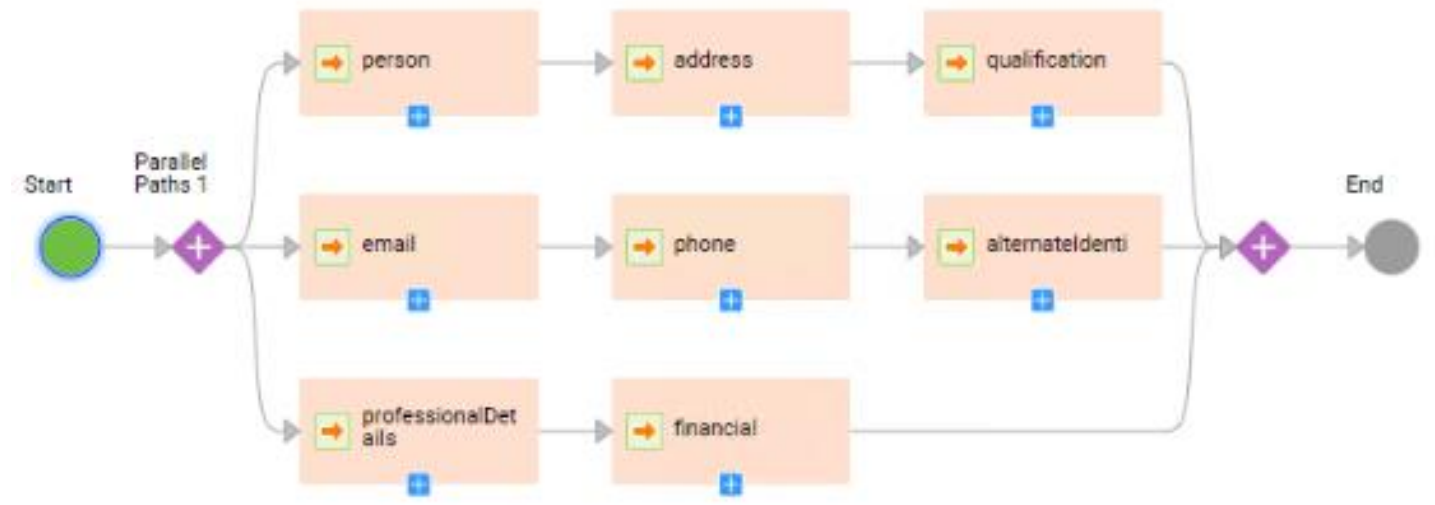
Scanner Update - Databricks

- Added support for scanning Databricks data assets using SQL Warehouse rather then all-purpose cluster

- For jdbc connection the scanner will use now "SQL Warehouse JDBC URL" provided in connection details

- Added support for proxy while scanning Databricks sources

Databricks
A cloud-based platform that combines Databricks Delta Lake and Databricks Notebooks.

Connection Details
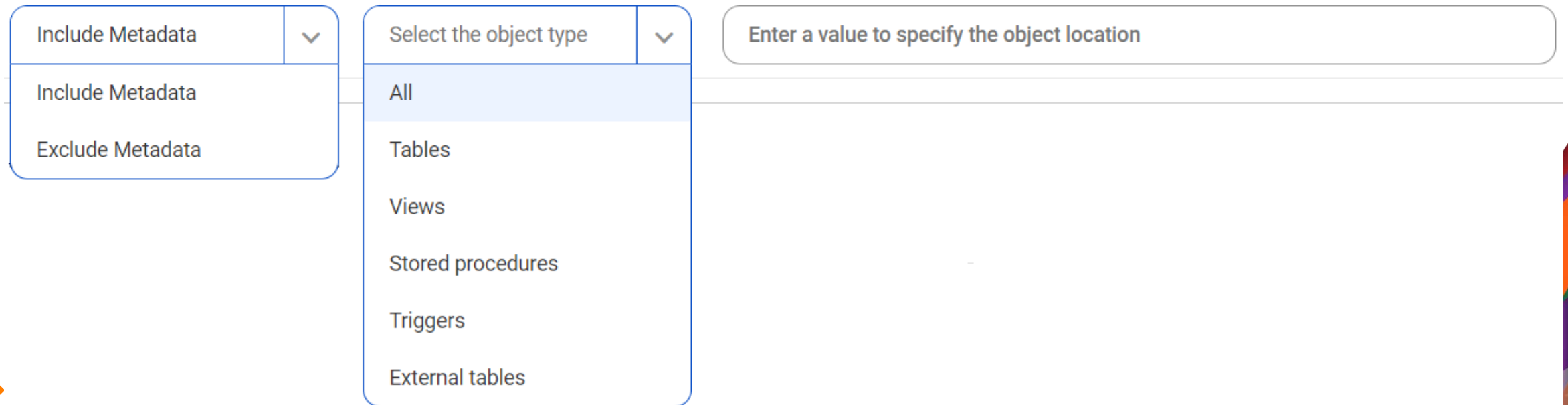
Connection Name:*	Databricks_SQLWarehouse_Conn
Description:	
Type:*	Databricks Delta
Runtime Environment:*	invrh79cdmp04
SQL Warehouse JDBC URL:	jdbc:databricks://partner-informatica-infarndawsd
Databricks Token:*	• • • • • • • •
Catalog Name:	

Connection Properties

Runtime Environment:	invrh79cdmp04
SQL Warehouse JDBC URL	jdbc:databricks://partner-informatica-━━━━━━:443/default;transportMode=http;ssl=1;httpPath=sql/protocolv1/o/(xxxxxxxxxxxxxxxx)/1003-113214-vmbksaed;AuthMech=3;UID=token;PWD=xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx;UseProxy=1;ProxyHost=xxxxxxxxxx;ProxyPort=3128
JDBC Driver Class Name	com.databricks.client.jdbc.Driver
Staging Environment	Personal Staging Location
S3 Authentication Mode	Permanent IAM Credentials
Use EC2 Role to Assume Role	false
S3 Region Name	US East (N. Virginia)
ADLS Endpoint	core.windows.net

Informatica

Scanner Update – Enhanced filtering

- There are introduced filtering enhancements for the following database scanners:

    - Oracle

    - Snowflake

    - Microsoft SQL Server

    - Azure Synapse

- Enhanced filtering capabilities to enable more precise analysis scope definition

Scanner Update – Promoted to GA

- Apache Atlas

  - added support for Spark jobs lineage extraction

  - added support for HBase data sources

  - added support for Sqoop process lineage with additional relational sources types like MySQL and PostgreSQL

- Microsoft Purview

  - added support for lineage extraction for database and file-based sources

  - added filtering on asset level

  - expanded list of supported asset types

- AWS Glue

  - improved support for scanning jobs executions with parameters

Apache Atlas
A scalable and extensible set of core foundational governance services

Microsoft Purview
A unified data governance, risk, and compliance solution

AWS Glue
A serverless ETL service from Amazon Web Services

Informatica

Scanner Update – HDInsight and EMR

Scanners certifications introduced in September and October patch releases

- Apache Hive
  You can extract metadata from Apache Hive source systems hosted on:

  - Amazon EMR cluster

  - Azure HDInsight

- Hadoop Distributed File System
  You can extract metadata from Hadoop Distributed File System source systems hosted on:

  - Amazon EMR cluster

  - Azure HDInsight

Cloud Data Quality

What is new for Cloud Data Quality and Profiling?

November 2023 Release Updates

- ❑ Profiling Features
    - ❑ Ability to download failed and valid rows from scorecard
- ❑ Profiling PAM Updates
    - ❑ AWS Athena
    - ❑ Databricks using IDMC connector in Advanced runtime
    - ❑ Generic JDBC
    - ❑ SAP BW, SAP ECC and SAP S/4 Hana as a source in a source mapplet.
    - ❑ First "N" Rows for ODBC in PostgreSQL and DB2

- ❑ Exception Management:
    - ❑ New email notification with task summary
    - ❑ Configure Advanced Rule for exception
- ❑ Ability to Convert existing Basic Rule logic into Rule Advanced Mode
- ❑ Verifier:
    - ❑ Added new output ports for USPS CASS Cycle O
    - ❑ Certified Reports Available
    - ❑ Updated to use latest 6.40 engine
- ❑ Updated Data Quality Financial Bundle
    - ❑ Added US Mortgage rule specifications

Informatica®

Cloud Data Profiling

What sources can you profile? (November 2023)

- 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 (October 2023)**
- Snowflake (using native V2 driver)
- Google Big Query
- Delimited File on: File system | Amazon S3 | Azure ADLS Gen2 | Google Cloud Storage
- ODBC
- **JDBC (November 2023)**
- SAP ECC, SAP S/4 Hana and SAP BW

- **Databricks Delta tables using IDMC drivers (November 2023)**
  - **[Requires 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
  - Azure ADLS Gen2 -
  - **[Requires Advanced / Elastic]**

Deprecated
- Azure Synapse, Snowflake and Redshift using ODBC (still works, not recommended)
- Databricks Delta tables using ODBC driver

Cloud Data Profiling

Databricks using IDMC connector

- Problem statement:
  - Profiling Databricks only available via Databricks ODBC driver

- Solution:
  - Added support for Databricks using IDMC driver
  - Supports
    - Managed and External Tables in Delta Format
    - Unity Catalog
    - Complex Data types

- Requires Advanced (Elastic) runtime
  - Mandatory to support complex data types

Cloud Data Profiling

Ability to download failed and valid rows from scorecard

- Problem statement:

  - User want to share some rows with data quality issues (Invalid Rows) in scorecard

- Solution:

  - Ability to download the "Failed" and "valid" rows from UI

  - Format:

    - Delimited CSV file

    - Has legend file to describe rule output columns

  - Requires "Drill Down" permission

Preview of Failed Rows

null_check(state): 25 Rows

city	postalcode	state	ssn	company	jobtitle	gender	churnflag
OAKLAND	94660	California	###-##-3547	Flashdog	Graphic Designer	Female	1
PITTSBURGH	15240	Pennsylvania	###-##-8424	Topiclounge	Staff Accountant IV	Female	1
BREA	92822	California	###-##-9517	Realfire	Nurse Practicioner	Male	1
HENRICO	23228	\N	###-##-0429	Skidoo	Statistician I	Male	0
SAINT LOUIS	63167	Missouri	###-##-5335	Blogspan	General Manager	Male	1
SPRINGFIELD	62794	Illinois	###-##-6510	Riffwire	Systems Administr…	Male	0
DURHAM	27710	North Carolina	###-##-8636	Feedfire	Information Syste…	Male	1
PUEBLO	81005	Colorado	###-##-1103	Flipstorm	Structural Engineer	Male	1
TRENTON	8638	New Jersey	###-##-3429	Eire	Human Resources …	Male	1
PORT SAINT LUCIE	34985	Florida	###-##-5271	Plajo	Analog Circuit Desi…	Female	1
CLEVELAND	44105	Ohio	###-##-4034	Mudo	Senior Financial An…	Female	1

1 - 25 of 100 Items   Page 1 of 4   Items Per Page: 25

Close

Informatica

Cloud Data Quality

Ability to Convert existing Basic Rule logic into Rule Advanced Mode

- Problem statement:
  - Customer already has rules in basic mode, but want to expand using advanced mode

- Solution:
  - Ability to convert basic rules into advanced rules

- Design:
  - One time change
  - Cannot convert back to basic

Exception Management

Configure Advanced Rule for exception

Rule Editor   Manage Inputs and Outputs   ✔

```
1     IF (Email = NULL and Phone = NULL AND SSN = NULL) THEN
2         Output = 'InValid'
3         ExceptionPriority = 'Critical'
4         ExceptionDescription = 'Whole record is null, critical exception'
5     ELSEIF Email = NULL THEN
6         Output = 'InValid'
7         ExceptionPriority = 'Critical'
8         ExceptionDescription = 'Email cannot be null'
9     ELSEIF (Phone = NULL) THEN
10        Output = 'InValid'
11        ExceptionPriority = 'Major'
12        ExceptionDescription = 'Phone is  null'
13    ELSEIF (SSN = NULL) THEN
14        Output = 'InValid'
15        ExceptionPriority = 'Major'
16        ExceptionDescription = 'SSN cannot be  null'
17    ENDIF
```

Exception Management

New email notification with Job summary

❑ **Revised / richer email update**

    ❑ Provides Job summary
       ❑ Source rows processed
       ❑ Exception rows identified
       ❑ Number of issues found

    ❑ Download Exception Data file link

    ❑ Top 10 Issues

    ❑ Top 10 rules identifying Issues

Job Summary

Job Summary: Exception task ran successfully: Identified 86.3% exceptions

Exception Task Job:	Exception_Multiple_Rules_P0_43680-3
Start Date:	Aug 03, 2023, 1:02 PM
End Date:	Aug 03, 2023, 1:02 PM
Started By:	Hawkdqrel1
Status:	Success

Source Rows / Exception rows / Number of issues

Source Rows: 183
Indicates the number of rows in the data source that the exception task job read.

Exception Rows: 158
Indicates the number of rows in which the exception task job found data quality issues.

Number of Issues: 292
Indicates the number of data quality issues that the exception task job found in the exception rows.

Percentage of Exceptions Found: 86.3%

Download exception Data file

Download Exception Data: Exception_Multiple_Rules_P0_43680-3

Counts per priority:

Critical 3	Major 123	Minor 166

Top 10 issues

Top 10 Issues
Shows the top 10 issues, each with a count of one or more

Description:	Count
Countries - USP, GBR, UK is excluded	131
Rule_Escalations_Exception - exceeds 10	54
more then 10 escalations	54
overpaid	24
Name is null, so entry seems to be invalid	11
name is null	8
Country name cannot be empty	7
Invalid record or bad record, all the entries are null	3

Top 10 Rules identifying issues

Top 10 Rules identifying issues
Shows the top 10 rules that identify issues, each with a count of one or more

Description:	Count
Rule_List_of_Values_33767	138
Rule_String_Int_Float_Exception_40479	89
Rule_Escalation_Exception_2106	54

Verifier

Added new output ports for USPS CASS* Cycle O

❑ DPV Confirmation Enhanced

❑ DPV Footnote4

❑ DPV Footnote5

❑ DPV Drop Indicator

❑ LACSLink False Positive

❑ Non-Delivery Days Flag (NDD Flag)

❑ ZIP5 VALID FLAG

* Coding Accuracy Support System (CASS)

© Informatica. Proprietary and Confidential.

63

Verifier
Certified Reports

What's New in CDMP? November

*Kate Amory, Senior Product Manager*

What is new in Cloud Data Marketplace?

November 2023 Release Updates

- My Data Page (restructure)

- Technical Information on Data Assets

- Delivery Management:
  - Delete unused delivery targets
  - De-select default delivery template
  - Ability to order without delivery set

- Usability Improvements:
  - Reduced space on Data Collection
  - Jumbo tooltip on Star Owner
  - User avatars on Ratings
  - Search Improvements
  - Direct Access Link on Orders
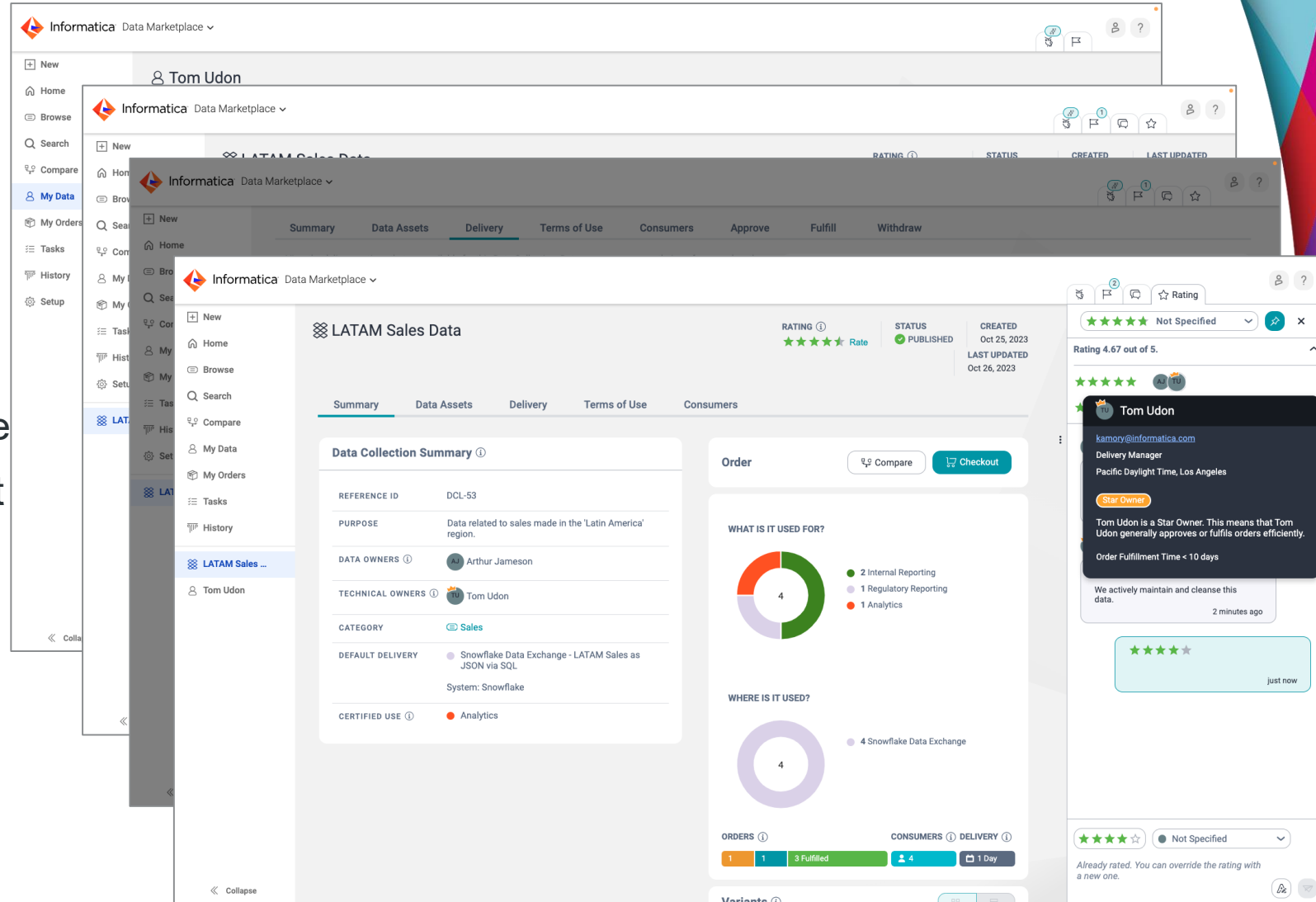
DEMO
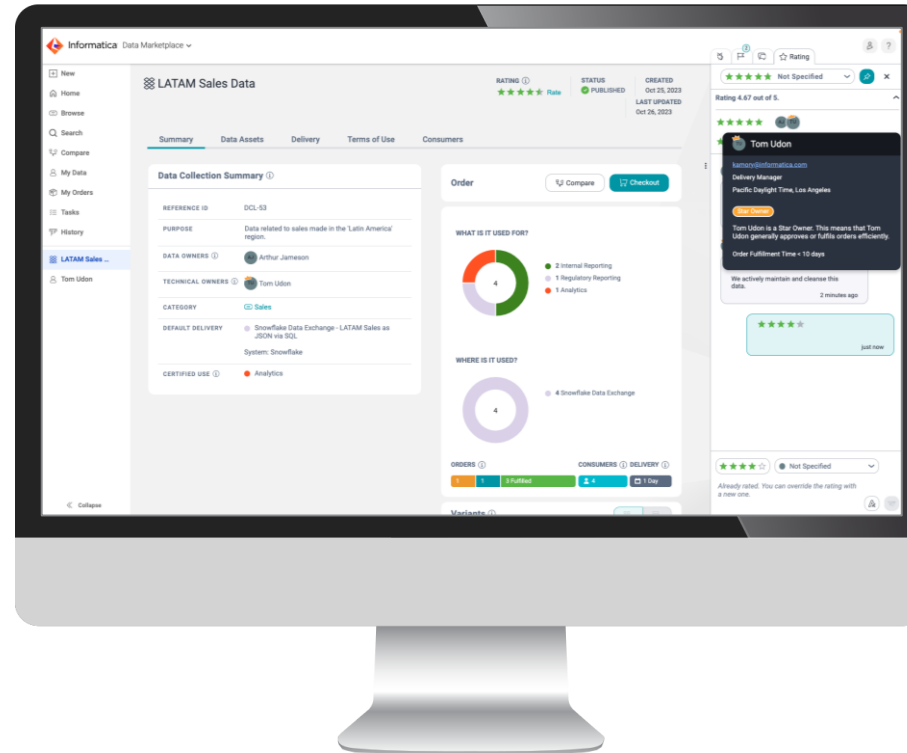
67

Amy Jenkins

Data User

Tom Udon

Delivery Owner

Thank You

Where data
& AI come to LIFE