

EDC Custom Scanners and IDQ Workflow Automation

CST K Rakesh Varma & Sachin Jain

30 Aug 2021



Agenda

- Overview
- Requirement
- Design
- Solution Scope and Automation Flow
- Demo
- Q&A

Overview

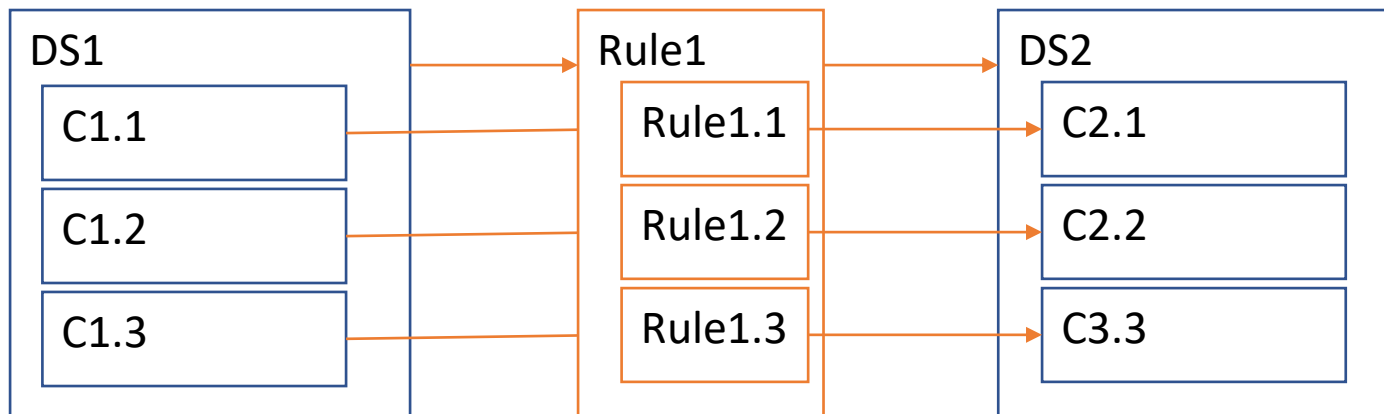
- Created a framework which will allow EDC users to extract data from comma separated file,
- And created EDC input files that can be loaded for Custom Metadata Scanner and Lineage Scanner
- Products Involved are IDQ and EDC REST API
- In EDC Once created a Custom Model, Resource type which can be used for similar Resource types
- And the framework will create input files to extract Customer Metadata and Lineage into EDC

Requirement

- EDC Customers who wants to reduce manual efforts in creating input file for EDC Custom Scanner both Metadata, Lineage files
- Can extend this framework and to meet custom file creation for different data model's
- Remove manual interventions in creating file which at times might induce the errors

Design

Workaround to display lineage link description



Rules will be provided along with sources/targets



lineage.csv



Objects.csv



Links.csv



Lineage1.csv

Rules to be added as description or custom objects



Informatica™

How this Solution Works

- Customer must follow same input file format every time to update the lineage from input file to EDC
- Convert transformation logic Space in the String to “_”(underscore)
- Convert empty/null value for String into “NoTransformation”
- Please make sure no special characters in the transformation logic string other than arithmetic operators
- Input file have to be csv file type

Automation Flow

- Run the IDQ workflow through a shell script
- This will create below files
 - objects.csv
 - links.csv
 - lineage_output.csv
 - objects.zip
- It will also upload below files into EDC so that you can schedule the workflow to pickup the latest files uploaded
 - objects.zip
 - lineage_output.csv
- Once resource's run in below you should be able to see the lineage extracted into EDC
 - Run your Transformation Custom Metadata Scanner
 - After successful completion run Lineage Custom Scanner

Sample File for Reference

- Github Link / Repository

https://github.com/infacst/INFACST_DOC

KB Article Reference for more detail on Creation Model / Resource Type

<https://network.informatica.com/docs/DOC-18212>

- Creation of Custom Metadata Resource

<https://docs.informatica.com/data-catalog/enterprise-data-catalog/h2l/1433-ingest-metadata-from-custom-resources/abstract.html>



DEMO

A nighttime cityscape featuring several illuminated skyscrapers. In the foreground, there are vibrant, multi-colored light trails (pink, purple, blue, and red) that suggest long-exposure photography of traffic. The sky is a deep blue. The overall composition is dynamic and modern.

Q&A
Thank You!