



CONFIDENTIAL

SAP Data Services

Document Version: 4.2 (14.2.0) – 2019-07-22

Match Blueprints User's Guide

Content

- 1 Introduction. 3**
- 1.1 Documentation set for SAP Data Services content objects. 3
- 1.2 SAP information resources. 3
- 1.3 Introduction to SAP Data Services 4.2 content objects. 4
- 2 Match Blueprints Overview. 6**
- 3 Downloading Blueprint Packages. 7**
- 3.1 Blueprint versions. 7
- 3.2 Available Match blueprints. 7
- 3.3 Downloading and setting up blueprints. 7
- 4 Configuring and Running Jobs. 10**
- 4.1 Editing the datastore 10
 - Microsoft SQL Server 10
 - Other database types. 11
- 4.2 Before running the jobs. 13
- 5 Match Review Blueprints. 14**
- 5.1 Preparation. 15
 - Adding an SAP Information Steward connection. 15
 - Running the "prep" job. 15
 - Importing the template tables. 16
- 5.2 MatchReview1: Data Consolidation in Data Services. 16
 - Running the "before" job. 16
 - Creating a match review configuration. 17
 - Performing a match review. 18
 - Running the "after" job. 19
- 5.3 MatchReview2: Data Consolidation in Data Services and Information Steward. 19
 - Running the "before" job. 19
 - Creating a match review configuration. 20
 - Performing a match review. 22
 - Running the "after" job. 23

1 Introduction

1.1 Documentation set for SAP Data Services content objects


You should become familiar with all of the pieces of documentation that relate to the SAP Data Services blueprints and other content objects.

Document	What this document provides
<i>Content Objects Summary</i>	Lists all of the available blueprints and other content objects and the jobs and other objects that they contain.
<i>Content Objects What's New</i>	Highlights the new and enhanced blueprints and other content objects available for this release.
<i>Data Quality Management Custom Functions User's Guide</i>	Contains instructions for downloading and importing custom functions.
<i>Data Quality Management Match Blueprints User's Guide</i>	Contains a list of available Data Quality Management match blueprints and instructions for downloading, configuring, and running them.
<i>Data Quality Management Microservices Blueprints User's Guide</i>	Contains a list of available Data Quality Management microservices blueprints and instructions for downloading, configuring, and running them.
<i>Data Quality Management Product Blueprints User's Guide</i>	Contains a list of available Data Quality Management product blueprints and instructions for downloading, configuring, and running them.
<i>Data Quality Management Regional Blueprints User's Guide</i>	Contains a list of available Data Quality Management regional blueprints and instructions for downloading, configuring, and running them.

1.2 SAP information resources

A global network of SAP technology experts provides customer support, education, and consulting to ensure maximum information management benefit to your business.

Useful addresses at a glance:

Address	Content
Customer Support, Consulting, and Education services https://support.sap.com/ 	Information about SAP support programs, as well as links to technical articles, downloads, and online forums. Consulting services can provide you with information about how SAP can help maximize your information management investment. Education services can provide information about training options and modules. From traditional classroom learning to targeted e-learning seminars, SAP can offer a training package to suit your learning needs and preferred learning style.

Address	Content
Product documentation http://help.sap.com/bods/	SAP product documentation.
Product Availability Matrix (PAM) https://apps.support.sap.com/sap/support/pam	Get information about supported platforms for SAP Data Services. Use the search function to search for Data Services. Click the link for the version of Data Services you are searching for.
SAP Data Services Community http://scn.sap.com/community/data-services	Get online and timely information about SAP Data Services, including forums, tips and tricks, additional downloads, samples, and much more. All content is to and from the community, so feel free to join in and contact us if you have a submission.
Blueprints https://wiki.scn.sap.com/wiki/display/EIM/SAP+Data+Services+Blueprints	Blueprints for you to download and modify to fit your needs. Each blueprint contains the necessary SAP Data Services project, jobs, data flows, file formats, sample data, template tables, and custom functions to run the data flows in your environment with only a few modifications.

1.3 Introduction to SAP Data Services 4.2 content objects

Welcome to SAP Data Services 4.2 version 14.2.0 Content Objects.

Data Services overview

SAP Data Services delivers a single enterprise-class solution for data integration, data quality, and data profiling that allows you to integrate, transform, improve, and deliver trusted data to critical business processes. It provides one development UI, metadata repository, data connectivity layer, run-time environment, and management console—enabling IT organizations to lower total cost of ownership and accelerate time to value. With SAP Data Services, IT organizations can maximize operational efficiency with a single solution to improve data quality and gain access to heterogeneous sources and applications.

Data Services Content Objects overview

We've identified a number of common scenarios that you are likely to perform with SAP Data Services. For each scenario, we've included a blueprint that is already set up to solve the business problem in that scenario. Each blueprint contains the necessary project, jobs, data flows, file formats, sample data, template tables, and custom functions to run the data flows in your environment with only a few modifications.

You can download the blueprint packages from the SAP Community Wiki. On the website, we periodically post new and updated blueprints, custom functions, best practices, whitepapers, and other content. You can refer to this site frequently for updated content and use the forums to provide us with any questions or requests you may have.

Instructions for downloading and installing the content objects are also located on the SAP Community Wiki.

2 Match Blueprints Overview

The Match blueprints illustrate a number of common data quality match scenarios that you are likely to perform with SAP Data Services. For each scenario, we've included a blueprint that is already set up to solve the business problem in that scenario. Each blueprint contains the necessary project, jobs, data flows, file formats, sample data, template tables, and custom functions to run the jobs in your environment with only a few modifications.

3 Downloading Blueprint Packages

3.1 Blueprint versions

The following table shows the version of the Data Quality Management blueprints that can be used for SAP Data Services. The blueprint version is displayed on the SAP Data Services Blueprints page of the SAP Community Wiki.

SAP Data Services version	Blueprint version	Blueprints available
4.2	4.2	Data Quality Management
4.1.1	4.1.1	Data Quality Management
4.1	4.1	Data Quality Management
XI 4.0	XI 4.0	Data Quality
XI 3.2	XI 3.2	Data Quality
XI 3.1	XI 3.0	Data Quality
XI 3.0	XI 3.0	Data Quality

3.2 Available Match blueprints

Each blueprints package contains sample jobs configured to illustrate best practice settings for common Data Quality Management use cases involving the matching process.

To see the contents of each blueprint, including jobs and custom functions, see the *Content Objects Summary*.

Blueprint	Description
Data Quality Management Blueprints – Match	Contains miscellaneous jobs configured to illustrate best practice settings for specific Data Quality Management matching use cases.

3.3 Downloading and setting up blueprints

Context

These are the general steps for downloading and setting up Data Quality Management blueprint packages for SAP Data Services.

i Note

These jobs include Data Cleanse. You should install the SAP-supplied person and firm cleansing package PERSON_FIRM before you import the Match .atl file.

Procedure

1. To access the SAP Community Network website, go to <https://www.sdn.sap.com/irj/boc/blueprints> in your web browser.
2. Log into your account using your username and password, or create a new account.
3. Review the list of available blueprint packages and other content objects and their descriptions to decide which to download.
4. Select the blueprint package that you want to download. A new page opens.
5. Select *View Document*.
6. Save the .zip file to the Tutorial Files folder in your installed SAP Data Services path. By default, this folder is installed to \Program Files\SAP BusinessObjects\Data Services\Tutorial Files for 32-bit Windows and \Program Files (x86)\SAP BusinessObjects\Data Services\Tutorial Files for 64-bit Windows.
If you are running on UNIX, the Tutorial Files folder exists only on the Windows client workstation, and you should download the .zip file there.
7. In the Tutorial Files folder in Windows Explorer, right-click the .zip file and select to extract the compressed (zipped) folders to the Tutorial Files folder. For example, if you use WinZip for file compression, right-click the .zip file and select **WinZip > Extract to here**.
Extracting creates subfolders and places the files in the appropriate location. The .atl file is saved to the Data Quality Mgmt Samples folder, and the sample data files are saved to the Data Quality Mgmt Samples\Match folder.
8. In the Designer, import the dqm_blueprints_match.atl file. In the Passphrase window, enter **dqm_blueprints_match** and click *Import*. Click *OK* to close the warning window.
Importing the file adds a project called DataQualityMgmtBlueprintsMatch to your object library. The project contains jobs whose names begin with DqmBlueprintMatch and contain in their name the Data Quality Management use case that they illustrate. Each job contains a data flow with the same name as the job. The import also adds a datastore called DqmBlueprintsMatch to your object library, and file formats whose names begin with DqmMatchIn and DqmMatchOut for the sample input and output data.
9. If you are running on UNIX, copy the input files to the job server machine and create the same folder structure that is on the Windows client workstation.
10. If you imported the blueprint .atl files using a Data Services Designer on 32-bit Windows and use a job server on 64-bit Windows, then you must copy the blueprint files to the Data Services installation of the job server machine.

Related Information

[Editing the datastore \[page 10\]](#)

4 Configuring and Running Jobs

4.1 Editing the datastore

Context

After you download the blueprint .zip file to the appropriate folder, unzip it, and import the .atl file in the Designer, you must edit the DqmBlueprintsMatch datastore.

The database that you use for running the blueprints does not need to be the same database that is used for the SAP Data Services repository. It can be a locally installed database system or any shared database system that you have access to create tables in and read from those tables.

The DqmBlueprintMatch_MatchDNB, DqmBlueprintMatch_SuppressDMA, DqmBlueprintMatch_MatchReview1After, and DqmBlueprintMatch_MatchReview2After jobs write to flat files on your file system. However, the DqmBlueprintMatch_MatchDNBPrep, DqmBlueprintMatch_SuppressDMAPrep, and DqmBlueprintMatch_MatchReviewPrep, and DqmBlueprintMatch_MatchReview1Before, and and DqmBlueprintMatch_MatchReview2Before jobs create the following tables which must exist before you run the corresponding jobs:

- DQM_BLUEPRINTS_MATCH_DMA
- DQM_BLUEPRINTS_MATCH_DNB
- DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2

Related Information

[Microsoft SQL Server \[page 10\]](#)

[Other database types \[page 11\]](#)

4.1.1 Microsoft SQL Server

Context

If you have access to write and read data to tables in Microsoft SQL Server 2000, 2005, or 2008, complete the following steps.

Procedure

1. Select the *Datastores* tab of the Local Object Library, right-click the DqmBlueprintsMatch datastore and select *Edit*.
2. In the *Edit Datastore DqmBlueprintsMatch* window, enter your repository connection information in place of the four *CHANGE_THIS* values.
3. Click *OK*. If the window closes without an error message, then the database is successfully connected.
4. Before you run the DqmBlueprintMatch_MatchDNB, DqmBlueprintMatch_SuppressDMA, DqmBlueprintsMatch_MatchReview1Before, and DqmBlueprintsMatch_MatchReview2Before jobs that access a database, run the following jobs to create the tables:
 - DqmBlueprintMatch_MatchDNBPrep
 - DqmBlueprintMatch_SuppressDMAPrep
 - DqmBlueprintsMatch_MatchReviewPrep
5. Before you run the DqmBlueprintsMatch_MatchReview1After and DqmBlueprintsMatch_MatchReview2After jobs, you must run the following jobs and perform the expected remediation tasks in Information Steward Match Review:
 - DqmBlueprintsMatch_MatchReview1Before
 - DqmBlueprintsMatch_MatchReview2Before

Next Steps

After you edit the datastore, you should be able to run the blueprint jobs.

4.1.2 Other database types

Context

If you have access to write and read data to tables in another database system (other than Microsoft SQL Server), complete the following steps.

Procedure

1. Select the *Datastores* tab of the Local Object Library, expand the DqmBlueprintsMatch datastore, and expand the *Template Tables* subfolder.
2. Delete the template tables. Right-click the following tables, select *Delete*, and select *Yes* to confirm your selection.
 - DQM_BLUEPRINTS_MATCH_DMA
 - DQM_BLUEPRINTS_MATCH_DNB

- DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS
 - DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1
 - DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2
3. Expand the Tables folder and delete the following tables:
 - DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1_ACT
 - DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2_ACT
 4. Delete the datastore. Right-click DqmBlueprintsMatch, select *Delete*, and select *Yes* to confirm your selection.
 5. Create a new datastore with the same name as the one you just deleted. In the *Datastores* tab of the Local Object Library, right-click in the white space and select *New*. In the *Datastore* name field, enter the name **DqmBlueprintsMatch**. In the Database type field, select your database system. Complete the remaining fields with the connection information to the database that you have access to.
 6. Click *OK* to close the Create New Datastore window.
 7. Open the DqmBlueprintMatch_MatchDNBPrep data flow and delete the target. Then add a new template table with the same name by selecting the Template Table icon from the buttons on the right menu and clicking the data flow canvas. In the Create Template window, enter **DQM_BLUEPRINTS_MATCH_DNB** for the Template name, and select the DqmBlueprintsMatch datastore in the In datastore field. Click *OK* to close the Create Template window. Connect the last transform to the template table.
 8. Open the DqmBlueprintMatch_MatchReviewPrep data flow and delete the three targets. Then add three new template tables with the same names. Select the Template Table icon from the buttons on the right menu, and click the data flow canvas. In the Create Template window, enter the template table name from the table below, and select the DqmBlueprintsMatch datastore in the datastore field. Click *OK* to close the Create Template window. Repeat these steps to create three template tables, and connect each template table to the appropriate transform:

Template table	Transform
DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS	JobStatusTable
DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1	StagingTable1
DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2	StagingTable2

9. Before you run the DqmBlueprintsMatch_MatchReview1Before and DqmBlueprintsMatch_MatchReview2Before jobs that access a database, run the DqmBlueprintsMatch_MatchReviewPrep job to create the tables.
10. Before you run the DqmBlueprintsMatch_MatchReview1After and DqmBlueprintsMatch_MatchReview2After jobs, you must import tables that are created during the remediation process in Information Steward Match Review. In the Datastores tab of the Local Object Library, expand the DqmBlueprintsMatch datastore. Right-click Tables and select *Import By Name*. In the Name field, enter **DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1_ACT** and click *Import*. Repeat these steps to import a table named DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2_ACT.
11. Before you run the DqmBlueprintsMatch_MatchReview1After and DqmBlueprintsMatch_MatchReview2After jobs run following jobs and perform the expected remediation tasks in Information Steward Match Review.
 - DqmBlueprintsMatch_MatchReview1Before
 - DqmBlueprintsMatch_MatchReview2Before

Next Steps

After you edit the datastore, you should be able to run the blueprint jobs.

4.2 Before running the jobs

Prerequisites

Before you run your jobs, you must complete the following tasks:

1. Copy the address cleanse reference files.
 - The DqmBlueprintMatch_SuppressDMA jobs use only the U.S. address cleanse reference files.
 - While the DqmBlueprintMatch_MatchDNB jobs contain some records from countries other than the United States, for the purpose of the blueprint you may copy only the U.S. address reference files.
 - The DqmBlueprintMatch_MatchReview jobs contain records from Brazil, France, Germany, India, Mexico, and the United States.
 - The DqmBlueprintMatch_NPerFirm and all four of the DqmBlueprintsMatch_Addr* jobs include only records from the United States.
2. Install the SAP-supplied person and firm cleansing package PERSON_FIRM before running jobs that include data cleansing.
3. Set accurate values in the substitution parameter configuration Configuration1:
 - `$$RefFilesAddressCleanse`—Enter the path location where you copied the address cleanse reference files.

5 Match Review Blueprints

The match review blueprints are a set of blueprints that illustrate how to use SAP Data Services and SAP Information Steward Match Review together to complete the cycle of identifying matching records, manually reviewing suspect match groups, and creating best records.

- **MatchReview1:** This set of blueprints illustrates creating the best record in Data Services after the suspect match groups are reviewed in Information Steward Match Review.
- **MatchReview2:** This set of blueprints illustrates creating the best record automatically on high-confidence match groups in Data Services and manually creating the best record on suspect match groups in Information Steward Match Review in conjunction with reviewing the suspect match groups.

The match review blueprints include the following jobs:

Job	Description
DqmBlueprintMatch_MatchReview1After	After review and remediation of suspect match groups formed in the MatchReview1Before job in Information Steward Match Review, this job creates best records by performing data consolidation within match groups.
DqmBlueprintMatch_MatchReview1Before	Identifies matching records that have similar firm and address data, and routes results to a staging table for the purpose of review and remediation of the suspect matches through SAP Information Steward Match Review.
DqmBlueprintMatch_MatchReview2After	Joins the staging table that contains the best records and non-matching records from the MatchReview2Before job, and the staging table that contains the results of the review and remediation, along with best record creation, through SAP Information Steward Match Review.
DqmBlueprintMatch_MatchReview2Before	Identifies matching records that have similar firm and address data, and separates high-confidence match groups and suspect match groups. High-confidence match groups are routed to receive best record creation by performing data consolidation within match groups, and the best records together with non-matching records are stored in a staging table. Suspect match groups are routed to a staging table for the purpose of review and remediation, along with best record creation, through SAP Information Steward Match Review.
DqmBlueprintMatch_MatchReviewPrep	Creates staging tables and a job status table, without inserting any data, in preparation for running the DqmBlueprintMatch_MatchReview1Before and BlueprintMatch_MatchReview2Before jobs. Note that this is simply a preliminary step to running the subsequent blueprints, since the best practice in production is not to create these tables with SAP Data Services.

The following sections describe how to configure and run the match review blueprints.

5.1 Preparation

5.1.1 Adding an SAP Information Steward connection

Context

Log into the Central Management Console (CMC) and add a new SAP Information Steward connection to the database that has the staging tables.

Procedure

1. In your browser, log into the CMC as an administrator.
2. From the *Organize* area, select *Information Steward*.
3. In the left pane, select *Connections*.
4. In the open white space in the right pane, right-click and select **► New > Connection ▾**. Complete the Create Connection window with the following information:

Option	Value
Connection Name	DQM_Blueprints
Connection Type	Database Connection
Purpose	For data review

5. Complete Database Type, Server Name, Database Name, User Name, and Password with the same connection information that is used in the Data Services datastore DqmBlueprintsMatch.
6. Click *Save*.

5.1.2 Running the "prep" job

Context

In Data Services Designer, run the *DqmBlueprintMatch_MatchReviewPrep* job.

This job creates the job status table and the staging tables for both the MatchReview1 and MatchReview2 jobs.

5.1.3 Importing the template tables

Context

Procedure

1. In Data Services Designer, click the *Datastores* tab of the local object library.
2. Expand the *DqmBlueprintsMatch* datastore, and then expand *Template Tables*.
3. Right-click *DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1* and select *Import Table*.
The *DQM_BLUEPRINTS_MATCH_REVEIW_STAGING1* table is removed from the *Template Tables* group and now appears in the *Tables* group.
4. Right-click *DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2* and select *Import Table*.
The *DQM_BLUEPRINTS_MATCH_REVEIW_STAGING2* table is removed from the *Template Tables* group and now appears in the *Tables* group.

5.2 MatchReview1: Data Consolidation in Data Services

Scenario: Engage data stewards and domain experts in the manual review and correction of all match groups, but perform best record creation and consolidation logic in SAP Data Services.

There are two blueprints for this scenario. The "before" blueprint demonstrates how to populate match results in staging tables for match review. The "after" blueprint demonstrates how to consume match review results.

5.2.1 Running the "before" job

Context

In Data Services Designer, run the *DqmBlueprintMatch_MatchReview1Before* job.

This job outputs all data (high-confidence match groups, suspect match groups, and non-matching records) to the *DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1* staging table, and inserts one row into the *DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS* job status table to indicate that the job is pending match review.

5.2.2 Creating a match review configuration

Context

Procedure

1. Log into SAP Information Steward.
2. In the *Manage* list, select *Match Review Configurations*.
3. Click the *New* button to create a new match review configuration.
4. Complete Step 1 of the Create Match Review Configuration wizard to review only suspect match groups and to allow the reassignment of master record using the following values:

Option	Value
Name	Match Review 1
Connection	DQM_Blueprints
High Threshold	93
Approval Required After Review	Deselect
Master Record Can Be Reassigned	Select
Enable Best Record Creation	Deselect

5. Complete Step 2 of the Create Match Review Configuration wizard to read data from the staging table created by the "before" job. Click the *Select* button and select the *DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1* table. Then, in the *Columns to Display on the Match Review UI* section, click the *Select* button and select the following columns in this order:
 1. Country
 2. Organization
 3. Address
 4. City
 5. Region
 6. Postcode
6. Complete Step 3 of the Create Match Review Configuration wizard to read and write to the job status table. Click the *Select* button, select the *DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS* table, and select the *Indicate Status Change* option.
7. Complete Step 4 of the Create Match Review Configuration wizard to assign the administrator as the reviewer. Select the *Default User Assignment* row and click *Edit*. Select *Administrators* as a reviewer, click *OK*, and click *Finish*.

i Note

If you want to assign match review tasks to different users, make sure that all match review users are part of the predefined Data Review User user group and all of them have View permission for the DQM_Blueprints connection.

You can also distribute match groups among different users by defining multiple new user assignments. For more information, see the *SAP Information Steward Administrator Guide*.

8. In the Match Review Configurations window, select the *Match Review 1* row in the top pane and click *Run Now*. Wait until "Completed and created 1 task(s)" appears in the Last Run column. The task list displays the task as a row in the lower pane. The task name is the configuration name entered in Step 1 of the wizard with an underscore and the job run ID.
9. Click *Close*.
10. In the My Worklist tab of Information Steward, click the refresh button. The match review task is displayed for all reviewers and approvers.

5.2.3 Performing a match review

Context

Procedure

1. Click the match review task in the My Worklist tab and click *Open*.
2. Click the *Match Group List* tab on the left side.
The Match Group List displays one row per match group.

Each high-confidence match group displays "Auto Confirmed" in the Status column, and each suspect match group displays "For Review".
3. Click the *Filter by Status* combo box, deselect *Auto Confirmed*, and click *Apply*.
The Match Group List now displays only the suspect match groups.
4. In the Match Group List, select the first row and click *Open*.
5. Review each suspect match group. For each group that should remain matching, optionally reassign the master and click *Confirm*. For each suspect group that should not match, click *Unmatch* and click *Confirm*. After you confirm the last group, the Match Group List again displays and the status for each group is now "Reviewed".
6. Click *Close*.

Results

Match Review made the following changes to the tables completed by the "before" blueprint:

- DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS: The value in the JOB_STATUS column changed from "Pending Match Review" to "C" for this Job Run ID.
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1: No changes made to this table.
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1_ACT: Match Review created this activity table. After all match groups are reviewed in Match Review, the table contains one or more rows per match group. For

high-confidence match groups and for suspect match groups that received no change, one row was inserted for the master record of the match group. For match groups that received a reassignment of the master record, two rows were inserted to indicate the original master and the newly assigned master. For a two-record match group that was unmatched, the subordinate record was inserted with a 'U' as the new group rank. For a three-or-more-record match group that received an unmatched action, each subordinate record unmatched was inserted with a 'U' as the new group rank.

5.2.4 Running the "after" job

Context

In Data Services Designer, run the [DqmBlueprintMatch_MatchReview1After](#) job, which:

- Joins the data from the DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1 staging table and the DQM_BLUEPRINTS_MATCH_REVIEW_STAGING1_ACT activity table, including the changes that Match Review reported in the activity table.
- Performs data consolidation by pulling data from the subordinate records and overwriting the data in the master record of each match group.
- Removes the subordinates to output only the best records (consolidated master records and all non-matching records).

5.3 MatchReview2: Data Consolidation in Data Services and Information Steward

Scenario: Engage data stewards and domain experts in manual review of suspect match groups only. You can modify the suspect match groups, if necessary, and create the best record for the match groups.

There are two blueprints for this scenario. The "before" blueprint demonstrates how to populate match results in staging tables for match review. The "after" blueprint demonstrates how to consume match review results

5.3.1 Running the "before" job

Context

In Data Services Designer, run the [DqmBlueprintMatch_MatchReview2Before](#) job.

This job performs the matching process, identifies whether match groups are suspect or high-confidence, and then splits the data:

- The suspect match groups are output to the DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2 staging table.

- The high-confidence match groups receive data consolidation by pulling data from the subordinate records and overwriting the data in the master record of each match group. The subordinates are then removed to output only the best records (consolidated master records and all non-matching records).

The job inserts one row into the DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS job status table to indicate that the job is pending match review.

5.3.2 Creating a match review configuration

Context

Procedure

1. Log into SAP Information Steward.
2. In the *Manage* list, select *Match Review Configurations*.
3. Click the *New* button to create a new match review configuration.
4. Complete Step 1 of the Create Match Review Configuration wizard to review all match groups, and to create best records by updating the master records. using the following values:

Option	Value
Name	Match Review 2
Connection	DQM_Blueprints
Approval Required After Review	Deselect
Master Record Can Be Reassigned	Select
Define Best Record Creation Process	Update Master Record

5. Complete Step 2 of the Create Match Review Configuration wizard to read data from the staging table created by the "before" job and to select to display the columns that are to be used in the best record creation. Click the *Select* button and select the *DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2* table. Then, in the *Columns to Display on the Match Review UI* section, click the *Select* button and select the following columns in this order:
 1. Country
 2. Organization
 3. Address
 4. Address_Dual
 5. Subcity
 6. City
 7. Region
 8. Postcode
 9. Secondary_Address_Length

- 10. Contact_Name
- 11. Contact_Phone

6. Complete Step 3 of the Create Match Review Configuration wizard to define the best record creation rule for Organization and the group of address fields. To define the best record for Organization to update with the longest organization name, select the *Organization* row, click *Define*, complete the Define Best Record Strategy window as follows, and click *OK*.

Option	Value
When To Apply	Always
Main Strategy	String Length
Type	Longest String
1st Tie Breaker Strategy	Completeness

7. To define the best record for the group of address fields to update with the address data when one of the records has a secondary unit and the best record is missing one, select the *Address* row, click *Define*, complete the Define Best Record Strategy window as follows, and click *OK*.

Option	Value
When To Apply	Always
Main Strategy	Match Record Priority
Type	Highest Priority
Column	Secondary_Address_Length
1st Tie Breaker Strategy	Completeness

8. Repeat step 7 for the Address_Dual, Subcity, City, Region, and Postcode columns. Step 3 of the Create Match Review Configuration wizard should look like the following when completed:

Name	Main Strategy	1st Tie Breaker Strategy	2nd Tie Breaker Strategy	When to Apply
Country				
Organization	String Length	Completeness		Always
Address	Match Record Priority	Completeness		Always
Address_Dual	Match Record Priority	Completeness		Always
Subcity	Match Record Priority	Completeness		Always
City	Match Record Priority	Completeness		Always
Region	Match Record Priority	Completeness		Always
Postcode	Match Record Priority	Completeness		Always
Secondary_Address_Length				
Contact_Name				
Contact_Phone				

9. Complete Step 4 of the Create Match Review Configuration wizard to read and write to the job status table. Click the *Select* button, select the *DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS* table, and select the *Indicate Status Change* option.
10. Complete Step 5 of the Create Match Review Configuration wizard to assign the administrator as the reviewer. Select the *Default User Assignment* row and click *Edit*. Select *Administrators* as a reviewer, click *OK*, and click *Finish*.

i Note

If you want to assign match review tasks to different users, make sure that all match review users are part of the predefined Data Review User user group and all of them have View permission for the DQM_Blueprints connection.

You can also distribute match groups among different users by defining multiple new user assignments. For more information, see the *SAP Information Steward Administrator Guide*.

11. In the Match Review Configurations window, select the *Match Review 2* row in the top pane and click *Run Now*. Wait until "Completed and created 1 task(s)" appears in the Last Run column. The task list displays the task as a row in the lower pane. The task name is the configuration name entered in Step 1 of the wizard with an underscore and the job run ID.
12. Click *Close*.
13. In the My Worklist tab of Information Steward, click the refresh button. The match review task is displayed for all reviewers and approvers.

5.3.3 Performing a match review

Context

Procedure

1. Click the match review task in the My Worklist tab and click *Open*.
2. Click the *Match Group List* tab on the left side. The Match Group List displays one row per match group.
3. In the Match Group List select the first row and click *Open*.
4. Review the match group to determine whether it should be considered a match.
 - If the match group should remain matching, click *Next*.
 - If the match group does not match, click *Unmatch* and click *Confirm*.
5. For each match group that is determined to be a match, determine if data in the Contact_Name and Contact_Phone columns should be consolidated into the best record (the organization and address data is automatically consolidated). If one of the matching records in the lower pane has a phone number and the best record in the top pane does not, drag the data from the Contact_Phone column into the best record. If the records have different contact names and one of the matching records in the lower pane has a better name than the best record, then drag the data from Contact_Name into the best record. After you consolidate the data into the best record, click *Finish*.

The window displays the next match group.

6. Repeat steps 4 and 5 for each match group.

After you review the last match group, the Match Group List displays and the status for each group is now "Reviewed".

Results

Match Review made the following changes to the tables completed by the "before" blueprint:

- DQM_BLUEPRINTS_MATCH_REVIEW_JOB_STATUS: The value in the JOB_STATUS column changed from "Pending Match Review" to "C" for this Job Run ID.
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2: Master records were updated either when the best record definitions for organization or address columns moved data or when you manually updated the contact name or phone columns.
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2_ACT: Match Review created this activity table. After you review all match groups in Match Review, this table contains one or more rows per match group. For match groups that received no change, one row was inserted for the master record of the match group. For match groups that received a reassignment of the master record, two rows were inserted to indicate the original master and the newly assigned master. For a two-record match group that was unmatched, the subordinate record was inserted with a 'U' as the new group rank. For a three-or-more-record match group that received an unmatched action, each subordinate record unmatched was inserted with a 'U' as the new group rank.
- DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2_BR: Match Review created this best record lineage table. For each match group that data was moved either automatically with a best record definition or manually, a row for the master record and a row for each subordinate record that contained the source of the data were inserted into this table. In the subordinate record, a field value of "1" means that the field is where the moved data came from.

5.3.4 Running the "after" job

Context

In Data Services Designer, run the [DqmBlueprintMatch_MatchReview2After](#) job, which:



1. Joins the data from the DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2 staging table and the DQM_BLUEPRINTS_MATCH_REVIEW_STAGING2_ACT activity table and merges them.
2. Keeps the best records and inserts them into the output from the DqmBlueprintMatch_MatchReview2Before job that previously contained the best records from high-confidence match groups and all non-matching records.

Important Disclaimers and Legal Information

Hyperlinks

Some links are classified by an icon and/or a mouseover text. These links provide additional information.

About the icons:

- Links with the icon : You are entering a Web site that is not hosted by SAP. By using such links, you agree (unless expressly stated otherwise in your agreements with SAP) to this:
 - The content of the linked-to site is not SAP documentation. You may not infer any product claims against SAP based on this information.
 - SAP does not agree or disagree with the content on the linked-to site, nor does SAP warrant the availability and correctness. SAP shall not be liable for any damages caused by the use of such content unless damages have been caused by SAP's gross negligence or willful misconduct.
- Links with the icon : You are leaving the documentation for that particular SAP product or service and are entering a SAP-hosted Web site. By using such links, you agree that (unless expressly stated otherwise in your agreements with SAP) you may not infer any product claims against SAP based on this information.

Beta and Other Experimental Features

Experimental features are not part of the officially delivered scope that SAP guarantees for future releases. This means that experimental features may be changed by SAP at any time for any reason without notice. Experimental features are not for productive use. You may not demonstrate, test, examine, evaluate or otherwise use the experimental features in a live operating environment or with data that has not been sufficiently backed up.

The purpose of experimental features is to get feedback early on, allowing customers and partners to influence the future product accordingly. By providing your feedback (e.g. in the SAP Community), you accept that intellectual property rights of the contributions or derivative works shall remain the exclusive property of SAP.

Example Code

Any software coding and/or code snippets are examples. They are not for productive use. The example code is only intended to better explain and visualize the syntax and phrasing rules. SAP does not warrant the correctness and completeness of the example code. SAP shall not be liable for errors or damages caused by the use of example code unless damages have been caused by SAP's gross negligence or willful misconduct.

Gender-Related Language

We try not to use gender-specific word forms and formulations. As appropriate for context and readability, SAP may use masculine word forms to refer to all genders.

© 2019 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company. The information contained herein may be changed without prior notice.

Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies.

Please see <https://www.sap.com/about/legal/trademark.html> for additional trademark information and notices.