

Match Rule Configuration Example

Abstract

To match and merge duplicate records in the MDM Hub, you need to configure match rules. Use base object attributes to configure match rules. This article walks you through an example of how to configure match rules that the MDM Hub uses to match duplicate records in a base object.

Supported Versions

- Multidomain MDM 10.1 - 10.4

Table of Contents

Match Rule Configuration Example Overview.	2
Match Rule Configuration Scenario.	3
Configuring Match Rules.	4
Step 1. Review the Data.	4
Step 2. Identify the Base Objects for the Match Job.	5
Step 3. Configure Match Properties.	5
Configuring Match Properties.	5
Step 4. Define the Match Path.	6
Adding Match Path Components.	6
Step 5. Define Match Columns.	11
Defining Match Columns.	11
Step 6. Define a Match Rule Set.	17
Defining a Match Rule Set.	17
Step 7. Add Match Rules.	18
Adding Match Rules.	18
Step 8. Set Merge Options for Match Rules.	21
Setting a Match Rule as an Automerger Match Rule.	21
Step 9. Review the Match Properties.	22
Reviewing the Match Properties.	23
Step 10. Test the Match Rules.	24

Match Rule Configuration Example Overview

The match rules configuration example shows you how to configure match rules to match and merge duplicate records in the MDM Hub. The match rules configuration example is based on the data available in the MDM Hub sample Operational Reference Store (ORS) that is included in the MDM Hub Resource Kit.

To determine records that are duplicates, that is, matches of each other, the MDM Hub uses match rules. You configure match rules for records in base objects. You can define a match rule for exact or fuzzy matching. An exact match rule matches records that have identical values in match columns. A fuzzy match rule matches similar records based on probabilistic match determinations that consider likely variations in data patterns, such as misspellings, transpositions, omissions, and phonetic variations. The match rules that you configure depend on the characteristics of the data, and the particular match and merge requirements.

The match rules configuration example walks you through the configuration process in a step-by-step manner. The example is based on the Party base object data that is available in the MDM Hub sample ORS. The example shows you how to configure match rules for data of individuals in the Party base object to match and merge data based on names and addresses. To view the preconfigured match rules that are similar to the match rules described in the example, set up the MDM Hub sample ORS. For more information about the sample Operational Reference Store, see the *Informatica MDM Multidomain Edition Sample ORS Configuration Guide*.

Match Rule Configuration Scenario

Your organization has customer information stored in the Party base object and other related base objects. The Party base object contains duplicate records for many customers. You want to configure match rules that can identify and queue up duplicate records for merge.

The Party base object contains duplicate customer records that might have the first name and the last name missing, or with incorrect or inconsistent entries. The Display Name column has values for each record. The records are categorized as a Person or an Organization party type. The addresses of the customers are stored in the related Address base object. The Party Address Rel relationship base object defines the relationship between the records in the Party and the Address base objects.

You want to create match rules to match duplicate customer records that have the Person party type in the Party base object. Records with the Person party type belong to individuals. You can perform the match based on names and addresses stored in the Party and Address base objects.

To match duplicate records, you need to create an automerge match rule for records that are sufficiently similar and can be queued for automerge. Also, you need to create a manual merge match rule for records that are possible duplicates, but need to be reviewed by a data steward before the merge process.

The Party base object shows important columns and sample records.

Rowid Object	First Name	Last Name	Organization Name	Display Name	Party Type
1019	NULL	NULL	NULL	WILL R DE HAAN	Person
1072	NULL	NULL	NULL	AHMED RAUF	Person
1106	RACHEL	ARSEN	NULL	RACHEL ARSEN	Person
1154	RACHEL	ARSEN	NULL	RACHEL ARSEN	Person
1191	WILLIAM	DE HAAN	NULL	WILLIAM DE HAAN	Person
1419	BILL	DE HAAN	NULL	BILL ROGER DE HAAN	Person
1475	NULL	NULL	RYERSON AREA	RYERSON AREA MED CTR	Organization
1642	AHMED	RAUF	NULL	AHMED RAUF	Person
1800	NULL	NULL	NULL	RYERSON AREA MED CTR	Organization

The Address base object shows important columns and sample records.

Rowid Object	Address Line1	Address Line2	City Name	State Cd	Postal Cd
920	69 BUTLER ST	NULL	ATLANTA	NULL	30303
991	7610 ROSENWALD LN	NULL	NOKESVILLE	NULL	20181
1221	RR 1 BOX 4	NULL	SUGAR RUN	NULL	18846-9701
1279	RR 1 BOX 3	NULL	SUGAR RUN	NULL	18846-9701
1711	69 JESSE HILL JR DR SE	NULL	ATLANTA	NULL	30303-3033
1860	5493 S QUEENS RD	NULL	ROCHELLE	NULL	61068
1909	5193 S QUEENS RD	NULL	ROCHELLE	NULL	61068
1960	669 BUTLER ST SE	NULL	ATLANTA	NULL	30303-3033
2005	7601 ROSENWALD LN	NULL	NOKESVILLE	NULL	20181

The LU Address Type lookup base object shows important columns and sample address type lookups.

Rowid Object	Address Type	Address Type Disp	Address Type Desc
1	BILL	BILLING	BILLING
4	LGL	LEGAL	LEGAL

Rowid Object	Address Type	Address Type Disp	Address Type Desc
10	RSID	RESIDENCE	RESIDENCE
11	SHIP	SHIPPING	SHIPPING

The Party Address Rel relationship base object shows important columns and sample records.

Rowid Object	Party ID	Address ID	Address Type	Status CD
976	1019	920	BILL	NULL
1051	1072	991	BILL	NULL
1080	1191	1960	LGL	NULL
1100	1419	1711	BILL	NULL
2001	1154	1909	LGL	NULL
2002	1106	1860	LGL	NULL
2004	1642	2005	LGL	NULL
2049	1475	1279	LGL	NULL
2050	1800	1221	LGL	NULL

Configuring Match Rules

To configure match rules, perform the following tasks:

1. Review the data.
2. Identify the base objects for the match job.
3. Configure match properties.
4. Define the match path.
5. Define match columns.
6. Define a match rule set.
7. Add match rules.
8. Set merge options for match rules.
9. Review the match properties.
10. Test the match rules.

Step 1. Review the Data

Before you create the match rules, review and understand the data. The base objects might have incorrect, inconsistent, and missing values.

Review the customer data for quality of the data values, consistency, uniqueness, and logic. To create match rules to match individuals, you must understand the attributes related to individuals.

The sample data set includes the Party and Address base objects. The Party base object has the First Name and the Last Name columns with missing values. It is not ideal to base the match rules on the First Name and the Last Name columns, because of missing values. The Display Name column has values for all records and is good to use as a match column in match rules.

The Party Type column in the Party base object identifies a customer record as a person or an organization. You can use the Party Type column to filter out customer records that belong to organizations because you do not want to find matches for customer records that belong to organizations. You can improve match performance if you filter data that is not relevant for the match process.

The Address base object has columns such as Address Line1, City Name, and Postal Cd that you can use as match columns in match rules. The columns in the Address base object would help identify duplicate records in the Party base objects. You can create match rules to match individuals based on the name and address attributes.

Step 2. Identify the Base Objects for the Match Job

You want to match the data of individuals based on the name and address attributes. Before you create match rules for the data of individuals, you must identify the base objects that contain specific data on which you want to base the match rules.

In the example, the data of individuals is stored in the Party and Address base objects. The names of individuals are in the Party base object and the addresses of the individuals are in the Address base object. You can determine matches and eliminate duplicate records of individuals based on names and addresses in the Party and Address base objects. To find duplicates, you can run the match job on the Party base object that contains the names of individuals.

Step 3. Configure Match Properties

Before you configure the match columns for the match rules, configure the match properties for the Party base object, such as the match and search strategy. You want to configure the match properties for the Party base object because it has duplicate records for individuals.

You must decide on the match properties based on the characteristics of the data, your knowledge of the data, and the match and merge requirements. The data in the Party base object has imperfections such as spelling variations and transpositions. The fuzzy match and search strategy, which is based on probabilistic matching, is suitable for the Party base object data. When you set the fuzzy match and search strategy, you need to set a population for the match rule. The population defines the characteristics of the data that you want to use for matching. The data that you chose for matching is US-centric, and therefore you need to set the `US` population as a match property.

Configuring Match Properties

To configure the match properties, use the Schema tool in the Hub Console.

1. In the Hub Console, start the Schema tool.
2. Acquire a write lock.
3. In the navigation tree, expand the Party base object for which you want to configure match properties.
4. Click **Match/Merge Setup**.
The **Match/Merge Setup Details** page appears.
5. Click the **Properties** tab.
The **Party Match/Merge Setup Details** section appears.
6. To configure match rules, you must set the following properties:

Property	Value and Description
Match/Merge Strategy	Fuzzy. Probabilistic match strategy that takes into account spelling variations, possible misspellings, and transpositions that are present in the data of individuals.
Fuzzy Population	US. A population that defines certain characteristics about the records that you want to match.

The following image shows the **Properties** tab of the **Match/Merge Setup Details** page:

Match/Merge Setup Details	
Match Rule Sets	Primary key match rules
Properties	Match Columns
Party Match/Merge Setup Details	
Match Columns	0
Match Rule Sets	0
Match Rules in Active Set	0
Primary key match rules	0
Maximum matches for manual consolidation	1000
Number of rows per match job batch cycle	10
Accept All Unmatched Rows as Unique	No
Match/Search Strategy	Fuzzy
Fuzzy Population	US
Match Only Previous Rowid Objects	<input type="checkbox"/>
Match Only Once	<input type="checkbox"/>
Dynamic Match Analysis Threshold (0=disa...	0

Step 4. Define the Match Path

To configure match rules for related records, configure match path components. Related records might reside in one base object or in multiple base objects.

To configure match rules that involve child records, you need to add path components to the root base object. When you add components to the match path, you define the connection between related parent and child base objects.

Add the following components to the match path:

- Party. Base object that contains data of individuals and organizations.
- Address. Base object that contains addresses of customers that are either individuals, or organizations.
- LU Address Type. Look up base object that contains address type lookups.
- Party Address Rel. Relationship base object that defines the relationships between the records in the Party and Address base objects.

Adding Match Path Components

Match paths define the connection between parent and child base objects. To configure match rules that include parent and child records, add match path components to the root base object. Also, add filters to include or exclude data during the match process.

1. In the **Match/Merge Setup Details** page for the Party base object, click the **Paths** tab.

2. Add the Party Address Rel relationship base object as a match path component.
 - a. From the Path Components section, select the **Root for C_Party** match path component.

The following image shows the **Paths** tab of the **Match/Merge Setup Details** page:

The screenshot shows the 'Match/Merge Setup Details' dialog box with the 'Paths' tab selected. The 'Path Components' section contains a table with one row: 'Root for C_PARTY', 'N/A', 'Party', 'N/A', and 'N/A'. The 'Filters' section is empty.

Display name	Component Name	Table Name	Direction	Check Mis...
Root for C_PARTY	N/A	Party	N/A	N/A

Column	Operator	Values
--------	----------	--------

- b. In the **Path Components** section, click **Add**.

The following image shows the **Add Path Component** dialog box:

The **Add Path Component** dialog box appears.

The screenshot shows the 'Add Path Component' dialog box. The 'Identity' section has fields for 'Display name', 'Physical name' (C_MT_), and 'Check For Missing Children' (checked). The 'Constraints' section contains a table with one row: 'Party Cross-Reference', 'Parent-to-Child', and 'Rowid Object'.

Table	Direction	Foreign Key On
Party Cross-Reference	Parent-to-Child	Rowid Object

- c. In the **Display Name** field, enter Party Address Rel as the name for the match path component that you want to display in the Hub Console.

The **Physical Name** field is populated based on the display name for the match path component that you enter. Physical name is the name of the path component in the database.

The following image shows the **Add Path Component** dialog box:

The **Add Path Component** dialog box appears with the name fields populated.

Identity	
Display name	Party Address Rel
Physical name	C_MT_PARTY_ADDRESS_REL
Check For Missing Children	<input checked="" type="checkbox"/>

Constraints:		
Table	Direction	Foreign Key On
Party Address Rel	Parent-to-Child	Rowid Object

- d. Ensure that the **Allow missing child records** option is enabled.

The option to allow missing child records is enabled by default. When enabled, matching occurs between the parent base object records and their associated child base object records. Matching occurs even if the parent base object records do not have child records in the child base object for which the option is enabled.

Note: Match key generation depends on the configuration of the **Allow missing child records** option. If the **Allow missing child records** option is disabled, and if the parent record has no child records, match keys are not generated for the parent record.

- e. From the Constraints section, select a constraint, and click **OK**.

The Party Address Rel path component appears in the Path Components section.

The following image shows the **Paths** tab of the **Match/Merge Setup Details** page:

Match/Merge Setup Details

Primary key match rules | Match Key Distribution | Merge Settings

Properties | **Paths** | Match Columns | Match Rule Sets

Path Components

Display name	Component Name	Table Name	Direction	Check Mis...	
[-] Root for C_PARTY	N/A	Party	N/A	N/A	+
Party Address Rel	C_MT_PARTY_ADDRESS_REL	Party Address Rel	Parent-to-Child	Yes	-

Filters

Column	Operator	Values	
			+

3. Add the Address base object to the match path as a child of the Party Address Rel match path component.
 - a. From the Path Components section on the **Paths** tab, select the **Party Address Rel** match path component.
 - b. Click **Add**.
The **Add Path Component** dialog box appears.
 - c. In the **Display Name** field, enter the name that you want to display in the Hub Console for the Address base object.

- d. From the Constraints section, select a constraint, and click **OK**.

The Address path component appears in the Path Components section as a child of the Party Address Rel match path component. The direction of the relationship appears as Child-to-Parent.

The following image shows the **Paths** tab of the **Match/Merge Setup Details** page with path components:

Match/Merge Setup Details

Primary key match rules | Match Key Distribution | Merge Settings

Properties | **Paths** | Match Columns | Match Rule Sets

Path Components

Display name	Component Name	Table Name	Direction	Check Mis...
[-] Root for C_PARTY	N/A	Party	N/A	N/A
[-] Party Address Rel	C_MT_PARTY_ADDRESS_REL	Party Address Rel	Parent-to-Child	Yes
[-] Address	C_MT_ADDRESS	Address	Child-to-Parent	Yes

Filters

Column	Operator	Values
--------	----------	--------

mdm_sample ... admin ... 63M of 494M

4. To include individuals and exclude organizations for matching, configure a filter for the Party base object that has the Party Type column.
 - a. From the Path Components section, select the **Root for C_Party** match path component.
 - b. In the Filters section, click **Add**.
The **Add Filter** dialog box appears.
 - c. From the **Column** list, select **Party Type**.
 - d. From the **Operator** list, select **IN**.
 - e. Beside the **Values** field, click **Edit**.
The **Edit Values** dialog box appears.
 - f. Click **Add**.
The **Add Value** dialog box appears.
 - g. In the Value field, enter **Individual**, and click **OK**.
 - h. Click **OK**.

The filter for the Party Type column appears in the Filters section. The filter for the Party Type column ensures that records such as Ryerson Area Med Ctr that belong to the party type Organization are excluded for match key generation. Only records such as Rachel Arsen and Ahmed Rauf that belong to the Person party type are included for match key generation.

Step 5. Define Match Columns

After you configure the match path components, define the match columns that you want to use in the match rules. You can configure both fuzzy and exact match columns because you chose a fuzzy match and search strategy for the Party base object.

You want to create match rules to match duplicate records of individuals. The match rules must include columns with names of individuals and address data to match duplicate records.

To match duplicate records of individuals, define the following columns as match columns:

- Display Name
- Address Line 1
- Address Line 2
- City Name
- State Cd
- Postal Cd

Defining Match Columns

Define match columns for the Party base object.

1. In the **Match/Merge Setup Details** page for the Party base object, click the **Match Columns** tab.
2. Configure the fuzzy match key settings for the Party base object.
 - a. From the **Key Type** list, select **Person Name**.

The **Path Component** is set to **Root (Party)**, which is the path component for the **Person_Name** fuzzy match key that is added to the Match Columns section.

- b. From the **Key Width** list, select **Standard**.

The search range for which the MDM Hub generates match keys is set to the standard size.

The following image shows the **Match Columns** tab with the Person Name fuzzy match key configured:

Match/Merge Setup Details

Primary key match rules | Match Key Distribution | Merge Settings
 Properties | Paths | Match Columns | Match Rule Sets

Fuzzy Match Key

Key Type	Person Name
Key Width	Standard
Path Component	Root (Customer)

Match Columns

	Field Name	Column Type	Path Component	Source Table
	Person_Name	Fuzzy Match Key	Root	Party

Match Column Contents - Source Table: Customer

Available columns:

- Display Name
- First Name
- Last Name
- Middle Name
- Organization Name
- Party Type

Selected columns:

The MDM Hub will generate match keys based on the Person Name fuzzy match key that you configured.

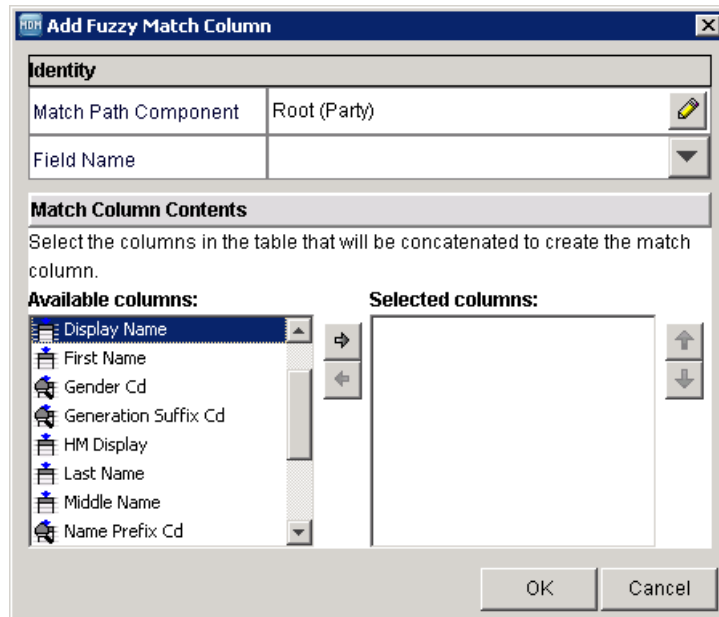
3. Define the Display Name column as the source column for the Person_Name fuzzy match key column.
 - a. In the Match Column Contents section, select the **Display Name** column from the **Available columns** list.
 - b. Click the right arrow.

The Display Name column moves to the **Selected columns** list, and the source column for the Person_Name match column is defined.

4. Add the fuzzy match column, Address_Part1, to contain street addresses.
 - a. Click **Add Fuzzy Match Column**.

The **Add Fuzzy Match Column** dialog box appears.

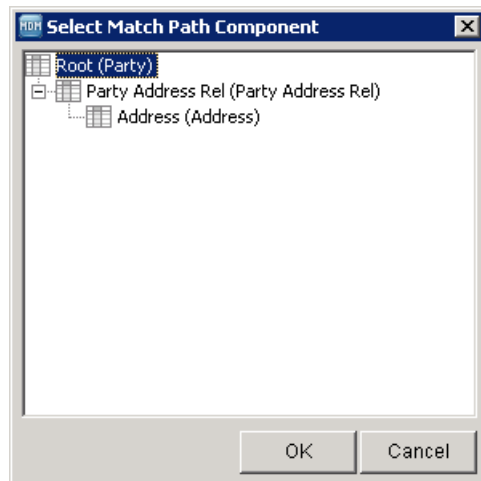
The following image shows the **Add Fuzzy Match Column** dialog box:



- b. Beside the **Match Path Component** field, click **Edit**.

The **Match Path Component** dialog box appears.

The following image shows the **Select Match Path Component** dialog box:



- c. Select the Address base object that contains the customer address data, and click **OK**.
 - d. From the **Field Name** list, select **Address_Part1**.
 - e. From the **Available columns** list, select the **Address Line 1** and **Address Line 2** columns that you want to concatenate to create the Address_Part1 match column.
 - f. To move the columns to the **Selected columns** list, click the right arrow.

The Address Line 1 and Address Line 2 columns move to the **Selected columns** list. The columns are concatenated to create the Address_Part1 match column. The **Add Fuzzy Match Column** dialog box appears.

The following image shows the **Add Fuzzy Match Column** dialog box:

Add Fuzzy Match Column

Identity

Match Path Component	Address (Address)	
Field Name	Address_Part1	

Match Column Contents

Select the columns in the table that will be concatenated to create the match column.

Available columns:

- Address Line3
- Bldg Name
- Care Of Address
- City Name
- Country Code
- County Name
- Dept Name
- Floor Number

Selected columns:

- Address Line1
- Address Line2

OK Cancel

- g. Click **OK**.

The Address_Part1 match column is added to the Match Columns section.

The following image shows the **Match Columns** tab with the match columns configured for person name and address data:

The screenshot shows the 'Match/Merge Setup Details' dialog box with the 'Match Columns' tab selected. The 'Fuzzy Match Key' section is at the top, followed by the 'Match Columns' table. Below that is the 'Match Column Contents - Source Table: Address' section, which includes 'Available columns' and 'Selected columns' lists.

Match/Merge Setup Details				
Primary key match rules		Match Key Distribution	Merge Settings	
Properties	Paths	Match Columns	Match Rule Sets	
Fuzzy Match Key				
Key Type	Person Name			
Key Width	Standard			
Path Component	Root (Party)			
Match Columns				
Field Name	Column Type	Path Component	Source Table	
Address_Part1	Fuzzy	Address	Address	
Person_Name	Fuzzy Match Key	Root	Party	
Match Column Contents - Source Table: Address				
Available columns:				
City Name				
Country Code				
County Name				
Dept Name				
Floor Number				
Is Valid Ind				
Latitude				
Selected columns:				
			Address Line1	
			Address Line2	

5. Add the fuzzy match column, Address_Part2, for customer address data to contain city names and the state codes.

- a. Click **Add Fuzzy Match Column**.

The **Add Fuzzy Match Column** dialog box appears.

- b. Beside the **Match Path Component** field, click **Edit**.

The **Match Path Component** dialog box appears.

- c. Select the Address base object that contains the customer address data, and click **OK**.

- d. From the **Field Name** list, select **Address_Part2**.

- e. From the **Available columns** list, select the **City Name** and **State Cd** columns that you want to concatenate to create the Address_Part2 match column.

- f. To move the columns to the **Selected columns** list, click the right arrow.

The City Name and State Cd columns move to the **Selected columns** list. The columns are concatenated to create the Address_Part2 match column.

- g. Click **OK**.

The Address_Part2 match column is added to the Match Columns section.

6. Add the exact match column, Postal_Area, to contain postal codes for customer addresses.

- a. Click **Add Exact Match Column**.

The **Add Exact Match Column** dialog box appears.

- b. Beside the **Match Path Component** field, click **Edit**.

The **Select Match Path Component** dialog box appears.

- c. Select the Address base object that contains the customer address data, and click **OK**.

- d. From the **Field Name** list, select **Postal_Area**.

- e. In the **Available columns** field name, select the **Postal Cd** column that you want to use to create the Postal_Area match column.

- f. To move the column to the **Selected columns** list, click the right arrow.

The Postal Cd column moves to the **Selected columns** list.

- g. Click **OK**.

The Postal_Area match column is added to the Match Columns section.

The following image shows the **Match Columns** tab with the match columns configured for person name and address data:

The screenshot shows the 'Match/Merge Setup Details' dialog box with the 'Match Columns' tab selected. The 'Fuzzy Match Key' section is at the top, followed by the 'Match Columns' table. Below the table is the 'Match Column Contents - Source Table: Address' section, which includes 'Available columns' and 'Selected columns' lists.

Match/Merge Setup Details				
Primary key match rules		Match Key Distribution	Merge Settings	
Properties	Paths	Match Columns	Match Rule Sets	
Fuzzy Match Key				
Key Type	Person Name			
Key Width	Standard			
Path Component	Root (Customer)			
Match Columns				
	Field Name	Column Type	Path Compon...	Source Table
	Address_Part1	Fuzzy	Address	Address
	Address_Part2	Fuzzy	Address	Address
	Person_Name	Fuzzy Match Key	Root	Party
	Postal_Area	Exact	Address	Address
Match Column Contents - Source Table: Address				
Available columns:				
Address Line3				
Bldg Name				
Care Of Address				
City Name				
Country Code				
Selected columns:				
Address Line1				
Address Line2				

Step 6. Define a Match Rule Set

You want to define a match rule set that has a search level set to Typical. You can create match rules in the match rule set that you define. You need to create at least one match rule set. Match rule sets can contain a logical set of match rules that have some common properties.

Defining a Match Rule Set

Define a rule set to which you can add match rules.

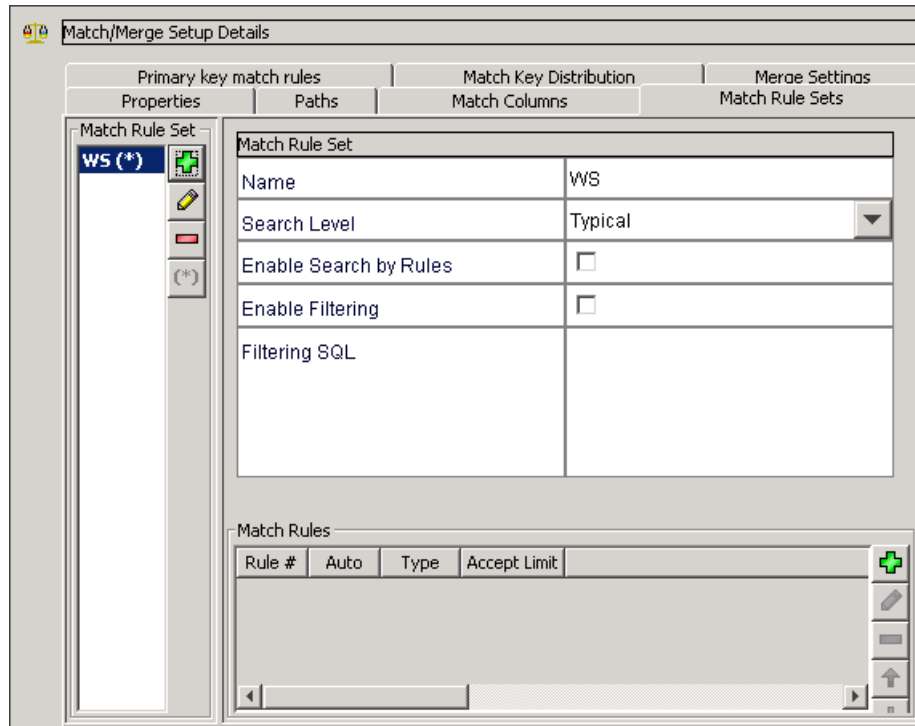
1. In the **Match/Merge Setup Details** page for the Party base object, click the **Match Rule Sets** tab.
2. To add a match rule set, click **Add**.

The **Add Match Rule Set** dialog box appears.

3. Enter a unique name, such as WS, for the rule set, and click **OK**.
4. From the **Search Level** list, select **Typical**.
The search level to search for match candidates is set to Typical.
5. To ensure that the match rule set is not reserved for exclusive use with the SearchMatch API, disable the **Enable Search by Rules** option.
6. Ensure that the **Enable Filtering** option is disabled.

Note: If the **Enable Filtering** option is disabled, all records are processed by the match rule set when a match batch job is run. If the **Enable Filtering** option is enabled, you can define a filter for this match rule set, and only the filtered records are processed by the match rule set.

The following image shows a Match/Merge Setup Details page with the WS match rule set configured:



Step 7. Add Match Rules

Before you add match rules, you must decide whether you want to set the match rules for automerge or manual merge. To match individuals at an address, you need a match purpose that has the name and address components.

If you set the match rules for automerge, the MDM Hub merges the records that are matched without the intervention of a data steward. If you configure the match rules for manual merge, the MDM Hub does not automatically merge the records that are matched. A data steward can review the matched records and start the merge job.

Also, you must decide whether you want to set up fuzzy, exact, or filtered match rules. A fuzzy match rule is required for data that might have inconsistencies. If the quality of the data is good, you can configure exact match rules. If you want to run a large batch job and at the same time ensure optimal performance, you might want to configure filtered match rules. Filtered match rules are exact match rules that use the fuzzy match key in addition to exact match columns.

To match individuals at an address, use the Resident match purpose that has the name and address components.

Adding Match Rules

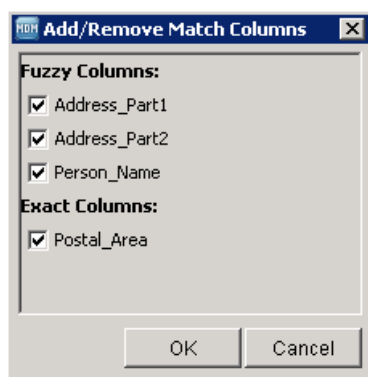
To match duplicate individuals based on name and address, add fuzzy match rules with the Resident match purpose to the WS match rule set. You created the WS match rule set in the preceding step.

1. In the **Match/Merge Setup Details** page for the Party base object, click the **Match Rule Sets** tab.
2. Add an manual merge match rule with the Typical match level.
 - a. In the Match Rules section, click **Add**.
The **Edit Match Rule** dialog box appears.
 - b. From the **Match/Search Strategy** list, select **Fuzzy**.

A fuzzy match/search strategy is a probabilistic match strategy that includes records with spelling variations, misspellings, and transpositions.

- c. Beside the **Match Columns** field, click **Edit**.
The **Add/Remove Match Columns** dialog box appears.
- d. Select the following columns:
 - Address_Part1
 - Address_Part2
 - Person_Name
 - Postal_Area

The following image shows the **Add/Remove Match Columns** dialog box with the match columns selected:



- e. Click **OK**.
The match column names appear in the **Match Columns** field.
- f. To identify matches for person names based on addresses, from the **Match Purpose** list, select **Resident**.
- g. To define the precision of the match between records, from the **Match Level** field, select **Typical**.
The Typical match level is suitable for most match jobs that match individuals. The Conservative match level results in fewer matches than the Typical match level and some potential matches might pass through the match process without being flagged as a match. The Loose match level might result in a significantly higher number of matches than the Typical match level, and might include matches that are not really matches.

The following image shows the **Edit Match Rule** dialog box with the fuzzy match rule settings:

Match Rule	
Match/Search Strategy	Fuzzy
Match Purpose	Resident
Match Level	Typical
Geocode Radius (in meters)	0
Accept Limit Adjustment	0

Match Columns	Match Properties
Address_Part1	Column Type:
Address_Part2	Match Subtype: <input type="checkbox"/>
Person_Name	Non-Equal Matching: <input type="checkbox"/>
Postal_Area	NULL Matching: <input checked="" type="radio"/> Disabled
	<input type="radio"/> NULL Matches NULL
	<input type="radio"/> NULL Matches non-NULL
	Segment Matching: <input type="checkbox"/>
	Segment Matches All Data: <input type="checkbox"/>
	Segment Match Values: <input type="text"/>

OK Cancel

- h. Click **OK**.

A manual merge match rule with the Resident match purpose appears in the Match Rules section.

3. Add a manual merge match rule with the Loose match level.

- a. In the Match Rules section, click **Add**.

The **Edit Match Rule** dialog box appears.

- b. From the **Match/Search Strategy** list, select **Fuzzy**.

A fuzzy match/search strategy is a probabilistic match that includes records with spelling variations, misspellings, and transpositions.

- c. Beside the **Match Columns** field, click **Edit**.

The **Add/Remove Match Columns** dialog box appears.

- d. Select the following fuzzy columns from the **Fuzzy Columns** options:

- Address_Part1
- Person_Name

- e. Click **OK**.

The fuzzy match columns appear in the **Match Columns** field.

- f. To identify matches for person names based on address, from the **Match Purpose** list, select **Resident**.

- g. To define the precision of the match between records, from the **Match Level** field, select **Loose**.

- h. Click **OK**.

A manual merge match rule with the Resident match purpose appears in the Match Rules section.

The following image shows the **Match Rule Sets** tab with two fuzzy, manual merge match rules configured for the WS match rule set:

Match/Merge Setup Details

Properties Paths Match Columns Match Rule Sets Primary key match rules Match Key Distribution Merge Settings

Match Rule Set

WS (*)

Match Rule Set

Name WS

Search Level Typical

Enable Search by Rules ☐

Enable Filtering ☐

Filtering SQL

Match Rules

Rule #	Auto	Type	Accept Limit	Purpose(Level)	Columns
1	No	Fuzzy	0	Resident(Typical)	Address_Part1 (Fuzzy) Address_Part2 (Fuzzy) Person_Name (Fuzzy) Postal_Area (Fuzzy)
2	No	Fuzzy	0	Resident(Loose)	Address_Part1 (Fuzzy) Person_Name (Fuzzy)

Auto

Step 8. Set Merge Options for Match Rules

During the match process, the match rule must determine whether to queue matched records for automerge or manual merge. You can toggle between manual merge and automerge for a match rule. To automerge records after the match process, you can set the typical match level for the match rule.

The precision of the match rule with the Resident match purpose and the typical match level lends itself for automerge.

Setting a Match Rule as an Automerge Match Rule

To queue matched records for automerge during the match process, set the match rule as an automerge match rule.

1. Select the manual merge match rule with the typical match level that you want to change to an automerge match rule.
2. Click **Move Up**.

The manual merge match rule changes to an automerge match rule.

The following image shows the **Match Rule Sets** tab with an automerge match rule and a manual merge match rule that is configured for the WS match rule set:

Match/Merge Setup Details

Properties Paths Match Columns Match Rule Sets Primary key match rules Match Key Distribution Merge Settings

Match Rule Set

WS (*)

Match Rule Set

Name WS

Search Level Typical

Enable Search by Rules ☐

Enable Filtering ☐

Filtering SQL

Match Rules

Rule #	Auto	Type	Accept Limit	Purpose(Level)	Columns
1	Yes	Fuzzy	0	Resident(Typical)	Address_Part1 (Fuzzy) Address_Part2 (Fuzzy) Person_Name (Fuzzy) Postal_Area (Fuzzy)
2	No	Fuzzy	0	Resident(Loose)	Address_Part1 (Fuzzy) Person_Name (Fuzzy)

3. Click **Save**.
The **Rule Set Evaluation** dialog box appears.
4. Click **OK**.
The match rules that you created are saved.

Step 9. Review the Match Properties

After you save the match rules configuration, review the match properties. The match properties summarize the match and merge setup.

The match and merge properties must show the following details:

- Number of match columns configured
- Number of match rule sets configured
- Number of match rules in the active match rule set
- Maximum matches for manual consolidation
- Number of rows for each match batch job cycle
- Type of match and search strategy
- Name of the fuzzy population

Reviewing the Match Properties

For a summary of the match rules setup, review the match properties.

- In the **Match/Merge Setup Details** page for the Party base object, click the **Properties** tab.

The **Party Match/Merge Setup Details** view appears.

The following table summarizes the match and merge setup:

Properties	Value
Match Columns	4
Match Rule Sets	1
Match Rules in Active Set	2
Primary key match rules	0
Maximum matches for manual consolidation	1000
Number of rows per match job batch cycle	10
Accept All Unmatched Rows as Unique	No
Match/Search Strategy	Fuzzy
Fuzzy Population	US
Match Only Previous Rowid Objects	Disabled
Match Only Once	Disabled
Dynamic Match Analysis Threshold (0=disabled)	0

The following image shows the **Properties** tab of the **Match/Merge Setup Details** page:

Match/Merge Setup Details	
Match Rule Sets	Primary key match rules
Properties	Paths
Match Columns	Match Key Distribution
Match Columns	Merge Settings
Match Columns	4
Match Rule Sets	1
Match Rules in Active Set	2
Primary key match rules	0
Maximum matches for manual consolidation	1000
Number of rows per match job batch cycle	10
Accept All Unmatched Rows as Unique	No
Match/Search Strategy	Fuzzy
Fuzzy Population	US
Match Only Previous Rowid Objects	<input type="checkbox"/>
Match Only Once	<input type="checkbox"/>
Dynamic Match Analysis Threshold (0=disa...	0

Step 10. Test the Match Rules

To test the match rules, run the match job and review the results.

You can test the match rules on a sample data set that is representative of the larger data set. To test the match rules, run the match job on the sample, representative data set. After the match job completes, review the match results.

You can view the match results in the C_PARTY_MTCH match table that is associated with the Party base object. Also, you can view the match results through Merge Manager in the Hub Console, and through the Potential Matches option in Informatica Data Director. Review the matched records to verify their accuracy.

The C_PARTY_MTCH match table shows important columns that contain the sample match result.

ROWID_OBJECT	ROWID_OBJECT_MATCHED	ROWID_MATCH_RULE	AUTOMERGE_IND
1191	1019	SVR1.JJ4J	0
1191	1419	SVR1.JJ4J	0
1154	1106	SVR1.JJ4E	1
1642	1072	SVR1.JJ4E	1

The ROWID_OBJECT column contains the row ID of the records that are matched to the records with row ID in the ROWID_OBJECT_MATCHED column.

The ROWID_MATCH_RULE column contains the row ID of the match rules that you create. The match rules are in the C_REPOS_MATCH_RULE table in the repository. SVR1.JJ4J is the row ID of the match rule with the Resident match purpose and a loose match level. SVR1.JJ4E is the row ID of the match rule with the Resident match purpose and a typical match level.

The example shows that the record for William De Haan that has row ID 1191 has two matches. The record matches the individual with the row ID 1019, Will R De Haan, and the individual with the row ID 1419, Bill Roger De Haan. The

matches are based on the name and address columns and are similar but need to be reviewed by a data steward before the merge. Both the matches of the record for William De Haan are queued for manual merge.

The match table shows that the record for Rachel Arsen that has row ID 1154 matches another record for Rachel Arsen that has row ID 1106. Both the records for Rachel Arsen have addresses in the Address base object that are similar and can be safely queued for automerge.

The following table shows the names of some individuals whose records were matched to other similar records:

Display Name	Display Name Matched
WILLIAM DE HAAN	WILL R DE HAAN
WILLIAM DE HAAN	BILL ROGER DE HAAN
RACHEL ARSEN	RACHEL ARSEN
AHMED RAUF	AHMED RAUF

The match result appears correct in the example. You can use the match rules that you configured for the larger match job.

Note: To view any issues that might occur during the match job, review the `cmxserver.log` file located in the following directory:

On UNIX. <infamdm_install_directory>/hub/server/logs

On Windows. <infamdm_install_directory>\hub\server\logs

Timeouts can occur because of match properties such as the number of rows for each match job batch cycle. You can change the properties in the **Properties** tab of the **Match/Merge Setup** page.

To check the progress of the match job and to get summary statics, review the Process Server log.

The Process Server log, `cmxserver.log`, is in the following directory:

On UNIX. <Process Server installation directory>/hub/cleanse/logs

On Windows. <Process Server installation directory>\hub\cleanse\logs

Author

MDM Documentation Team