CLEARMDM

clearMDM Data Stewardship Guide

Author: clearMDM
Document Version: v1.0
Product Version: v6.13
Date: 14th November 2022

This document provides an overview of the concepts and processes involved in the manual Data Stewardship of candidate matches.

Key Concepts

Data Stewardship

In the context of this document the terms Data Stewardship and Data Stewarding are equivalent in meaning and refer to the oversight and management of the MDM processes supported by clearMDM to ensure that candidate status record matches are correctly identified and stewarded efficiently through their lifecycle.

Match Record Pair Lifecycle

In technical terms, the clearMDM matching engine generates a **Matched Record Pair** record for every pair of records (e.g. Leads, Contacts and Accounts) that have matched at a score above the defined **Fuzzy Match Threshold Score %** (e.g. 75%) for the Target Object (e.g. Lead).

Where multiple records match together as a group (of transitive matches) the **Record Grouping Id** field (on the **Matched Record Pair** records) will be consistent across all pairs within the group.

Each **Matched Record Pair** describes the 2 records which have matched, the calculated match score % and threshold plus the match type (Fuzzy or Key). The **Match Status** field defines the current status of the match within the exclusive range of values Candidate, Accepted or Rejected. Where the **Match Score %** for a Matched Record Pair is above the **Auto Accept Match Threshold %** (as set for the Target Object) then the status value will be **Accepted** otherwise the status value will be **Candidate**. Alternatively, where Data Stewarding Conditions are implemented, if either record referenced by the matched record pair satisfies one or more condition, then the pair status will be Candidate irrespective of the match score (note, the Match Score % must be greater than the Fuzzy Match Threshold Score %).

The merge engine processes **Matched Record Pair** records by group and auto-merges groups which contain no Candidate status matched record pairs. Such groups are referred to as complete groups, as opposed to partial groups where data stewarding tasks remain. It's the task of the data steward to inspect **Candidate** status pairs and to set the status to **Accepted** or Rejected. Only when a group is fully stewarded will the merge process complete.

Match-Score-based Data Stewarding

The match-score data stewarding process is dynamically driven by the calculated match score strength versus the defined auto-accept threshold score. The intent of this method is to ensure that weak scoring matches (relative to the match score range) are inspected prior to auto-merge.

Condition-based Data Stewarding

The condition-based data stewarding process is dynamically driven by the attributes of the records that have matched. For example, it may be required that matches related to VIP Customers are data stewarded irrespective of the strength of the match score. The intent of this method is to provide explicit control over the records to be inspected prior to auto-merge.

Data Stewardship Processes
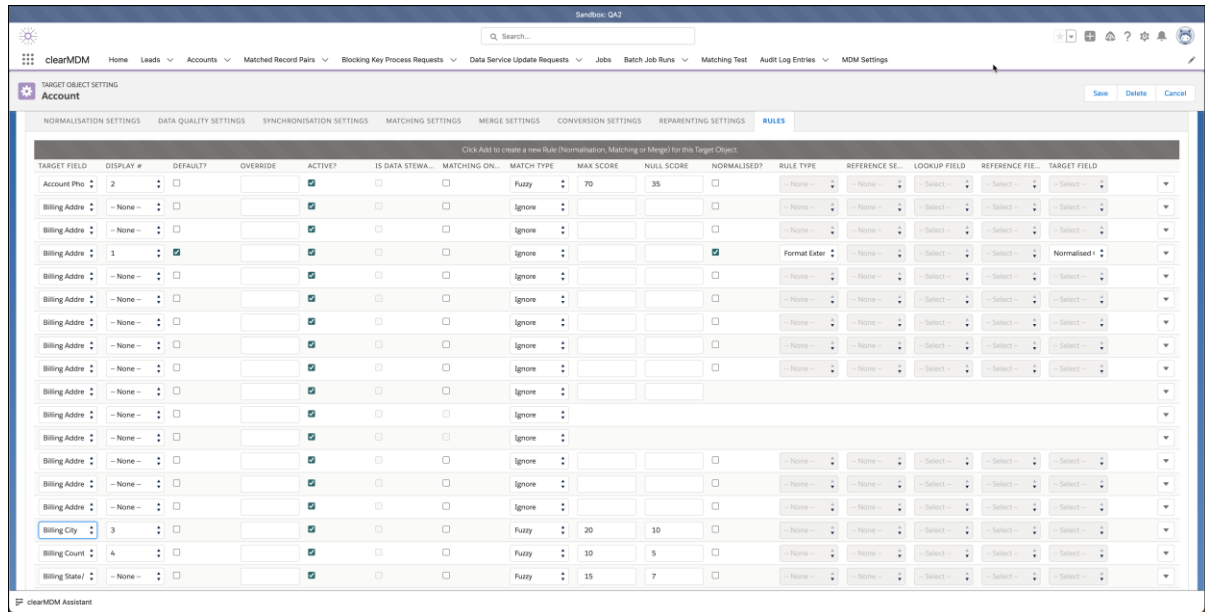
Define Match Score Thresholds

Screenshot 1 – Target Object Matching Setting

Screenshot 1 shows the MDM Settings UI page where the **Fuzzy Match Threshold Score %** and **Auto Accept Match Threshold %** are set.

This page is located at clearMDM App > MDM Settings tab > Target Objects subtab > [Select Target Object] > Matching Settings subtab.

Adjust the Matching Rules

Screenshot 2 – Target Object Matching Rules

Screenshot 2 shows the MDM Settings UI page where the **Matching Rules** are configured for a given **Target Object**. Please note, the term Target Object refers to the collective MDM configuration for a given Salesforce Object (e.g., Account, Contact or Lead) – or record type within that object – and is comprised of the settings and rules used for matching and merging data.

Matching Rule Types:

Ignore	Field is ignored for matching and active for merge only.
Fuzzy	• **Max Score**: Score value assigned when the 2 records being compared have an exactly equivalent field value, case insensitive. Otherwise, a % of the score value is assigned, where the % equates to the number of edits required to achieve equivalency of the field values as a % of the max possible edits. • **Null Score**: Score value assigned when either of the 2 records being compared has a blank field value.
Exact	• **Max Score**: Score value assigned when the 2 records being compared have an exactly equivalent field value, case insensitive. Otherwise, a zero score is assigned.
Deterministic	The 2 records being compared must have an exactly equivalent (or blank) field value, case insensitive, otherwise a zero % match score is assigned for the record-level comparison.

CLEARMDM

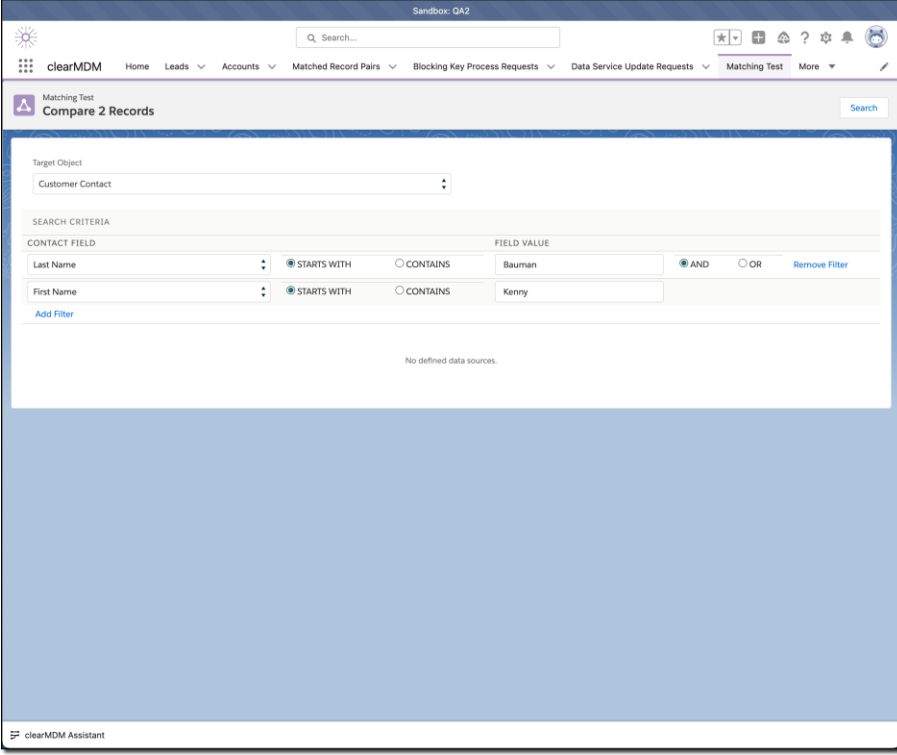| Deterministic Strict | The 2 records being compared must have an exactly equivalent (and non-blank) field value, case insensitive, otherwise a zero % match score is assigned for the record-level comparison. |
| Key | If the 2 records being compared have an exactly equivalent (and non-blank) field value, case insensitive, a 100% match score is assigned for the record-level comparison. |
| Key Approximate | • **Max Score**: Percentage of equivalency required. If the 2 records being compared have an approximately equivalent (and non-blank) field value, case insensitive, a 100% match score is assigned for the record-level comparison. |

Best Practice:

Key and Deterministic match rule types are processed in advance of fuzzy matching and provide a significant processing efficiency and performance benefit.

This page is located at clearMDM App > MDM Settings tab > Target Objects subtab > [Select Target Object] > Rules subtab.

Empirically Test the Match Score Calculation

clearMDM provides a **Matching Test** function which enables two records to be compared and a match score calculated; adjustments can also be made to the rules and the impact on the match score observed. The Matching Test features provides a convenient method to validate the matching rules against actual records at different anticipated levels of match scoring.

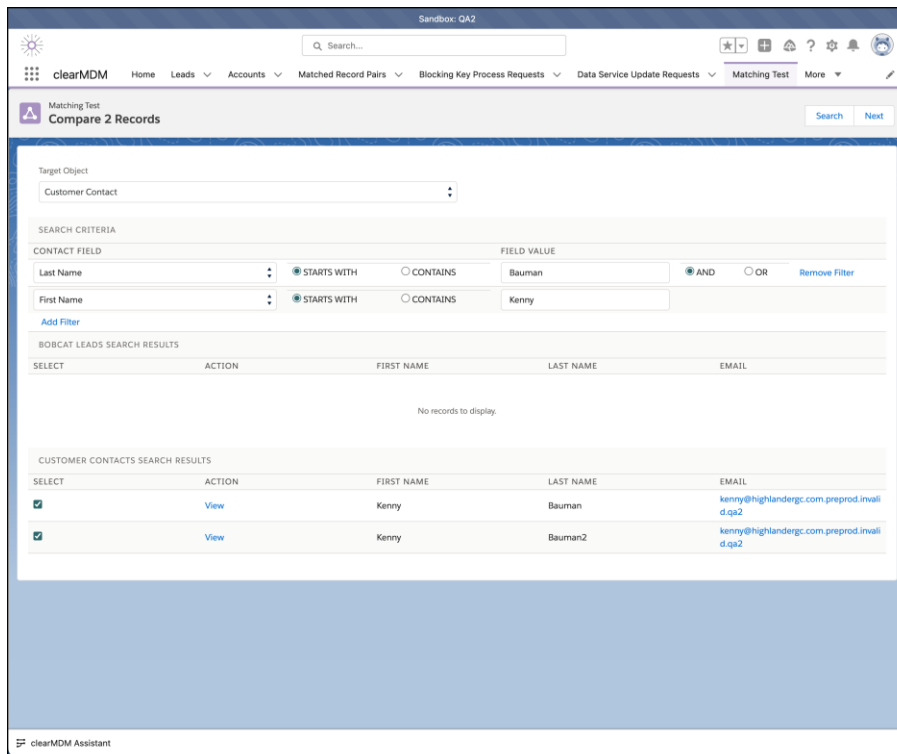*This page is located at clearMDM App > Matching Test tab.*

Screenshot 3.1 – Matching Test – Search for Records to Compare

The first step is to define a search filter to find the records to be compared. The Search button should then be clicked.

Matching Test – Select Records to Compare

Two records (only) within the search results should be selected and the Next button clicked.

Matching Test – Adjust Rules and Compare

The default view shows the list of fields active as matching rules and the record values.

Matching Test – View Outcome

Beneath the field list, the Compare Results panel shows the calculated match score and the effective thresholds. Changes can be made to the rule types and weightings with recalculation via the Compare button. The MDM Configuration can also be updated via the "Apply to Settings" button.

Identify Data Stewardship Tasks

Matched Record Pair records which are pending data stewardship (i.e. at Candidate status) can be found in the Matched Record Pairs tab within the MDM app.

Please note, not all Matched Record Pair records displayed in this tab will be Candidate status and require data stewarding. It is possible that Matched Record Groups will contain a mix of Candidate status and Accepted status matches; such groups are referred to as Partial Groups. The merge engine will skip partial groups until all matches (Matched Record Pairs) within the group have been assigned the Accepted or Rejected status.

Matching Record Pair Report

The Matched Record Pair data can be accessed via the Salesforce Reporting engine and used to build custom reports and dashboards.

Best Practice:

Salesforce Reporting Notifications allow a data steward user (or data stewarding team) to subscribe to a report which proactively notifies the list of data stewarding tasks for a given time period.

Action Data Stewardship Tasks

Data Stewarding tasks require the assignment of the **Accepted** or **Rejected** Match Status to **Candidate** Match Status **Matched Record Pair** records within a **Matched Record Group**. The objective should be to turn all partial groups to complete groups which contain only accepted or rejected pairs. Note, partial groups are ignored by automated record merge processes.

Match Status edits can be applied directly to the Matched Record Pair record, accessible from the Matched Record Pair tab, or from Matched Record Group page which can be accessed from the Matched Record Pair record. The Match Analysis page can be accessed from the same location and provides an overview of the match score calculation with a side-by-side field value view to assist with the data stewarding accept/reject decision-making process.

Matching Record Pair

Note, as Match Status edits are simply a field value update, such edits can be performed in bulk via List Views on the Matched Record Pair tab, or offline in a productivity tool such as Excel via file export/upload.

Match Analysis Page

The Matched Record Group page provides the complete view of all matches within a group and the current Match Status. Match Status values can be edited, and the group can also be merged automatically (Merge Now button) or manually Merged (Merge button).

Matched Record Group Page

The **Merge Now** button will execute merge in real-time and apply the full merge rule logic configured for the respective Target Object (e.g. Account, Contact or Lead). On completion, the Salesforce application will redirect to the master record for the group.

The **Merge** button initiates a manual merge process where the user can decide which field values are promoted to the master based on the data available on the records being merged. Note, the manual merge process ignores the current Match Status for all Matched Record Pairs in the group.

Best Practice:

Unless there is a specific reason to merge records manually, the Merge Now functionality should be utilised to merge records using the configured merge rules.

Alternatively, a record merge batch job can be scheduled to run periodically (hourly, nightly etc.) to process **Matched Record Groups** that have been completed by data stewarding activities.