CLEARMDM

QuickStart Guide 6 - Data Quality

Document Version: v1.4
Product Version: v3.13
Date: 16th November 2018

This document provides an overview and Step-by-Step implementation instructions for the clearMDM Data Quality MDM operation.

The document Appendices also provide additional reference materials.

For practitioner guidance in respect to the implementation of clearMDM please refer to the Implementation Model documentation provided on the website, or upon request.

Table of Contents

CLEARMDM

Data Quality Introduction

Definition

The Data Quality MDM operation establishes a clear understanding of data quality through Data Quality Rule Sets, configurable rules and statistic-driven actions. Proactive and actionable analysis of key data quality metrics can be imperative to the success of data-centric projects.

Standard Quality Rule Sets are installed by default (with accompanying analytics) and address common quality scenarios on the Account, Contact and Lead standard objects. Custom Quality Rule Sets can be defined to measure and enforce custom corporate data governance standards.

Key Concepts

Concept	Definition
Target Objects	A compatible object that is defined as the target for Data Quality operations and where Master Records will be created or updated. Account, Contact, Lead Standard Objects are typically configured as Target Objects, where duplicate records may exist directly in the object or indirectly in a separate object. A broad range of **Standard Objects** (including Person Accounts) are supported as Target Objects. **Custom Objects** are also supported.
Settings	Data Quality is enabled per Target Object on the Target Object Settings page. A setting-by-setting definition for the Data Quality Settings is provided in the appendices - Appendix A.
Rule Sets	A Data Quality Rule Set encapsulates a collection of quality rules that evaluate to a single Rule Set score value, plus a collection of quality actions that conditionally execute based on defined score value ranges. Rule Sets are processed in a defined order to allow the outputs of one Rule Set to be reference by subsequent Rule Sets. Rules Sets can be defined for Batch or Real-time evaluation or both. Batch evaluation requires the Data Quality MDM operation to be scheduled via the Jobs page, periodically or on an ad-hoc basis. Real-time evaluation occurs during record creation or update and enables quality scoring to be kept continually up to date.
Rules	Data Quality Rules express logical facts that evaluate to either True or False based on defined conditions (field population, related data, date offsets, format validation etc.).

	Where the fact is true, a defined score value is added to the evaluated Rule Set score. The score value assigned to each rule expresses the relative importance (or weighting) of that fact across the collection of rules defined within the Rule Set. New in v3.13: The score value can now not only be positive but also negative. So instead of adding score values together, it can be expressed where appropriate that the score value is reduced by a predefined amount when the criteria is met.
Actions	Data Quality Actions implement quality score-driven behaviour through mapping the evaluated Rule Set score to Field Update, Post-to-Chatter or Blocking actions. Actions can update a [Quality Score] field with the evaluated Rule Set score value, or block a record update where the quality score is decreased. Actions can also apply time-based conditions where action is required based on the passing of time, not on actual record modifications. New in v3.13: There is the ability to select from a score value of 0 to a predefined score.
Standard Rules and Actions	Standard Rule Sets are configured by default for the Account, Contact and Lead standard objects. clearMDM Mid Edition supports Standard Rule Sets only with the limitations below. Reference rules are not supported. Standard actions allow changes to the score value ranges only. No new actions can be added. Blocking actions are not supported. Enterprise Edition supports fully editable/extendable Standard Rule Sets and Custom Rule Sets.
Standard Analytics	The [clearMDM Data Quality Dashboards / Dashboard reports] folders contain a number of exemplar analytics built from the Standard Rules and Actions outputs.

Data Quality Use Cases

The table below outlines the primary use cases for the Data Quality MDM operation.

Use Case	Definition

CLEARMDM

Data Profiling	Multiple record-level quality statistics can be calculated by batch or real-time Rule Set processing and used as the basis for summary reporting that reveals data quality trends or patterns of issue. Profiling is typically the first stage of an iterative data quality management strategy.
Data Governance/ Compliance	Data Quality Rules and Actions enable the efficient measurement (and enforcement) of data governance standards.
Corrective Actions	Data Quality Rules Sets can identify invalid data conditions and apply corrective actions.
Preventative Actions	Data Quality Rules Sets can protect data quality by blocking record creation and modification operations that do not comply with the defined quality standards.
Time based Conditions	The Data Quality MDM operation running in Batch mode enables evaluation of date/datetime field values against set offsets from the current date. For example; the current age of a Customer can be evaluated using their Date of Birth and action taken when a significant age change occurs (Junior to Adult etc.). For example; travel booking records can be evaluated using the Travel Date field and action taken such as triggering itinerary sending when a specific offset of days (30 days before) is reached.
Lead or Case Scoring	Calculation of Lead/Case scoring based on Data Quality Rules and mapping of score value ranges to actions that update Rating and Priority values.
MDM: Master Record Priority	The Master Record Priority value is used by the Merge MDM operation when selecting the correct master record within a group. Often the selection logic must take into account custom business rules that include child record references. Quality Rule Sets can calculate the record-level **Master Record Priority** value using related data lookups often key to this logic, e.g. active product warranties, open cases etc.
MDM: Attribute Group Dynamic Priority	Attribute Groups allow groupings of fields to be merged collectively to the Master Record from the related Source Record with the highest priority. Quality Rule Sets can be used to calculate the

	quality scoring to be used with Attribute Groups of the dynamic priority type. For example; an Attribute Group may be defined for the Account Billing address which delegates priority to the Account Quality Score field. In such a model the Master Record Billing address fields would be populated from the Source Record with the highest quality address.
MDM: Data Source Partitioning	Automated, intelligent allocation of Source Records to partition data sources. Quality Rule Sets can calculate the record-level **Partition Value** used when allocating records to a partition Data Source. For example; high quality records can be allocated to the "High Quality Accounts" data source, whilst low quality records may be allocated to an "Inactive Accounts" data source.

CLEARMDM

Data Quality Methods

The table below outlines the supported methods for invocation of the Data Quality MDM Operation. Setting references refer to the Target Object Data Quality settings.

Method	Definition
Batch Job	The clearMDM Jobs page can be used to schedule the Data Quality MDM operation to run for a given Target Object immediately or on a scheduled basis. Where a daily batch processing model is implemented for MDM processing, the Data Quality MDM operation will typically be scheduled outside of the main processing sequence. Note, as all Target Object records are evaluated the processing time will reflect the data set size.
On Record Create	Records can be evaluated on creation to prevent creation of low quality records.
On Record Save	Records can be evaluated on save to prevent modifications that reduce the data quality scoring.
Action	Records can be evaluated by Process Builder as part of a custom process automation. To configure a Process Builder Action for this purpose add an Action with properties set as below. 1. Action Type = Apex 2. Apex Class = "Evaluate Records Action" 3. Record ID parameter = Reference [Object Id field]. The Action can also be implemented within Visual Workflow.
API	Records can be evaluated by a custom action exposed via the standard Force.com REST API; endpoint below. `/services/data/vXX.0/actions/custom/apex/clearmdm__RecordQualityAction` The API operation takes a single recordId parameter. Further details can be found in the clearMDM API Guide.

Step 1 – Configure Application Settings

Pre-requisite: clearMDM must be set to Active via the Application Settings page.
Pre-requisite: clearMDM must be set with Triggers Active via the Application Settings page.

CLEARMDM

To complete this step, first navigate to the MDM App, open the Settings tab, tick the two fields highlighted below and click the Save button.
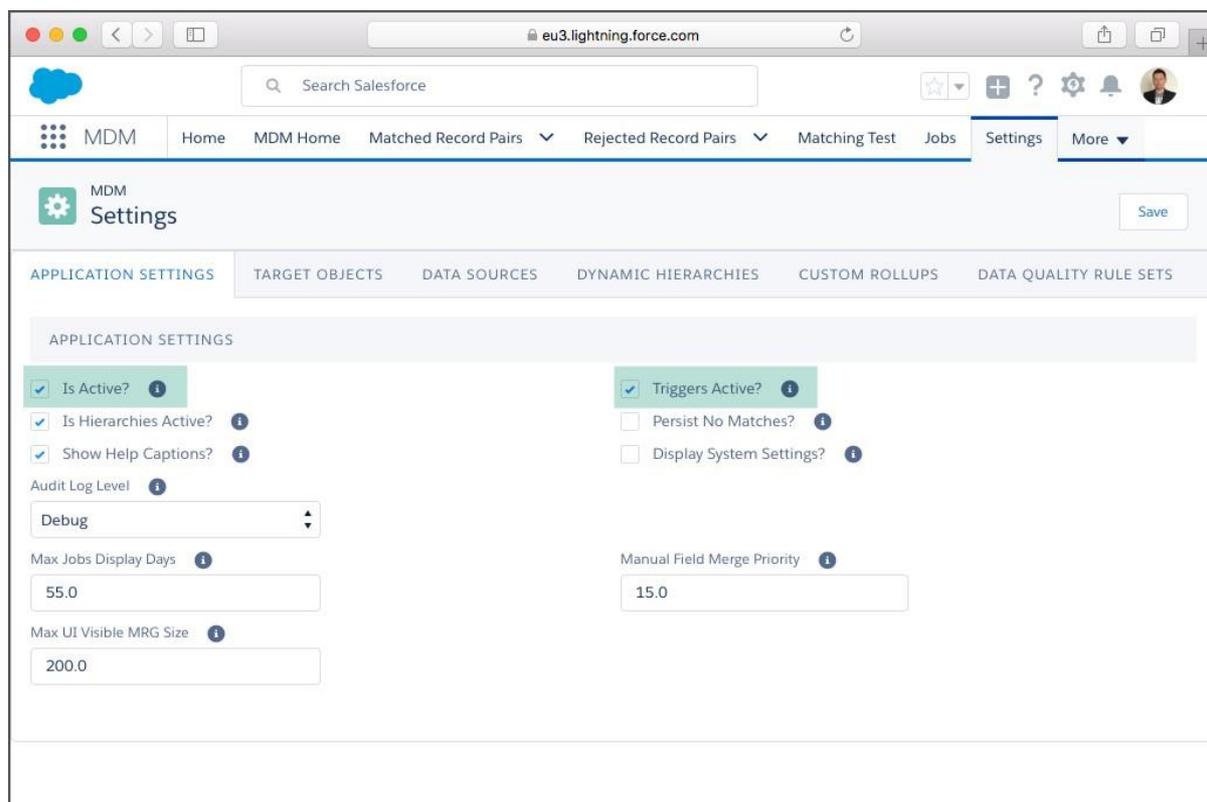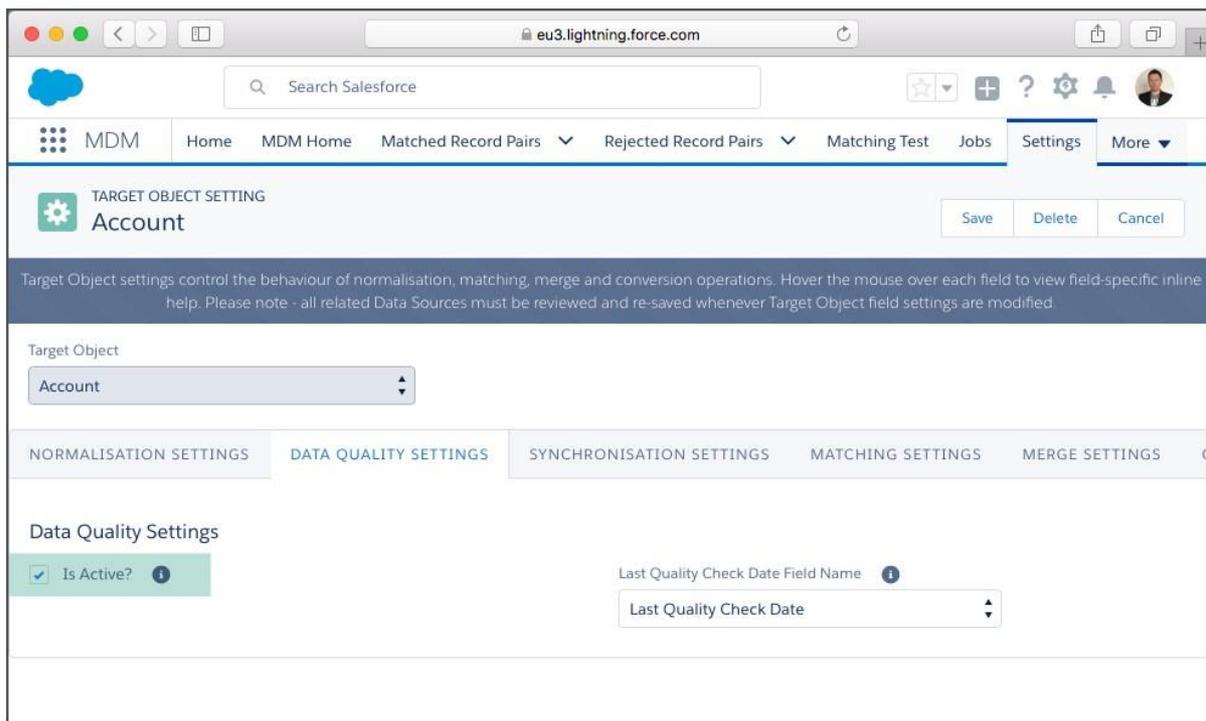

Step 2 – Activate the Target Object for Data Quality

To complete this step, first navigate to the MDM App, open the Settings tab, select the Target Objects tab and click the Edit link next to the required Target Object. Note, for Lightning Experience the Edit menu is accessible via the Dropdown menu in the rightmost table column.

Next, set the [Is Active?] flag equal to true in the Data Quality Settings section (or Data Quality Settings tab in Lightning Experience) and ensure the relevant settings are configured correctly.

Appendix A provides a Settings reference.

Step 3 – Configure Data Quality Rule Sets

To complete this step, first navigate to the MDM App, open the Settings tab, select the Data Quality Rule Sets tab and click the Edit link next to the required Data Quality Rule Set. Note, for Lightning Experience the Edit menu is accessible via the Dropdown menu in the rightmost table column.

Next, on the Data Quality Rule Set settings page make the required changes to the Rule Set, Rules and Actions.

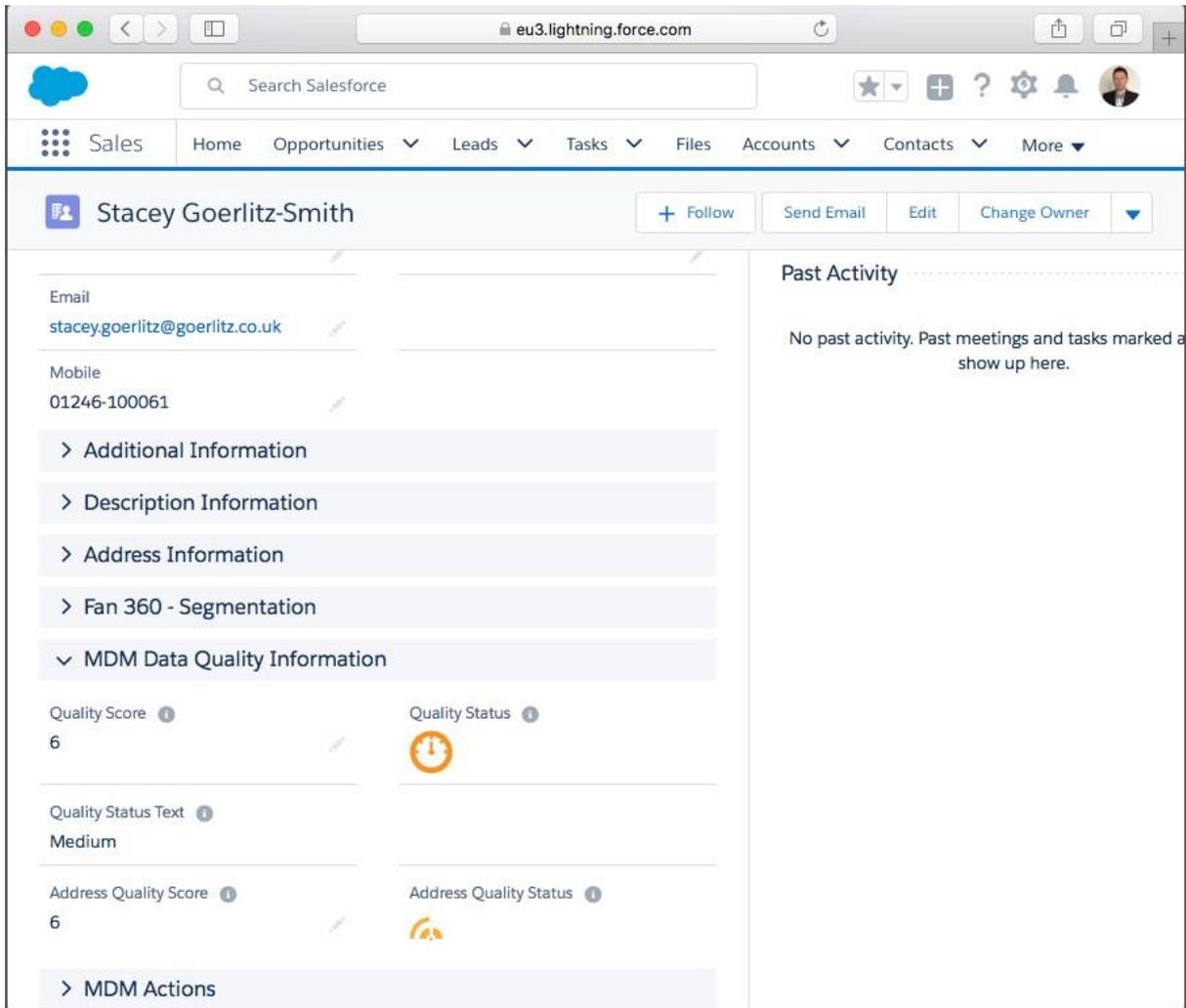Finally, click the Save button.

CLEARMDM

CLEARMDM

–

Step 4   Run the Data Quality Job

To complete this step, first navigate to the MDM App, open the Jobs tab, select the Reparenting Job type and the required Target Object. Finally select the required Job Schedule and click the "Schedule Job" button.
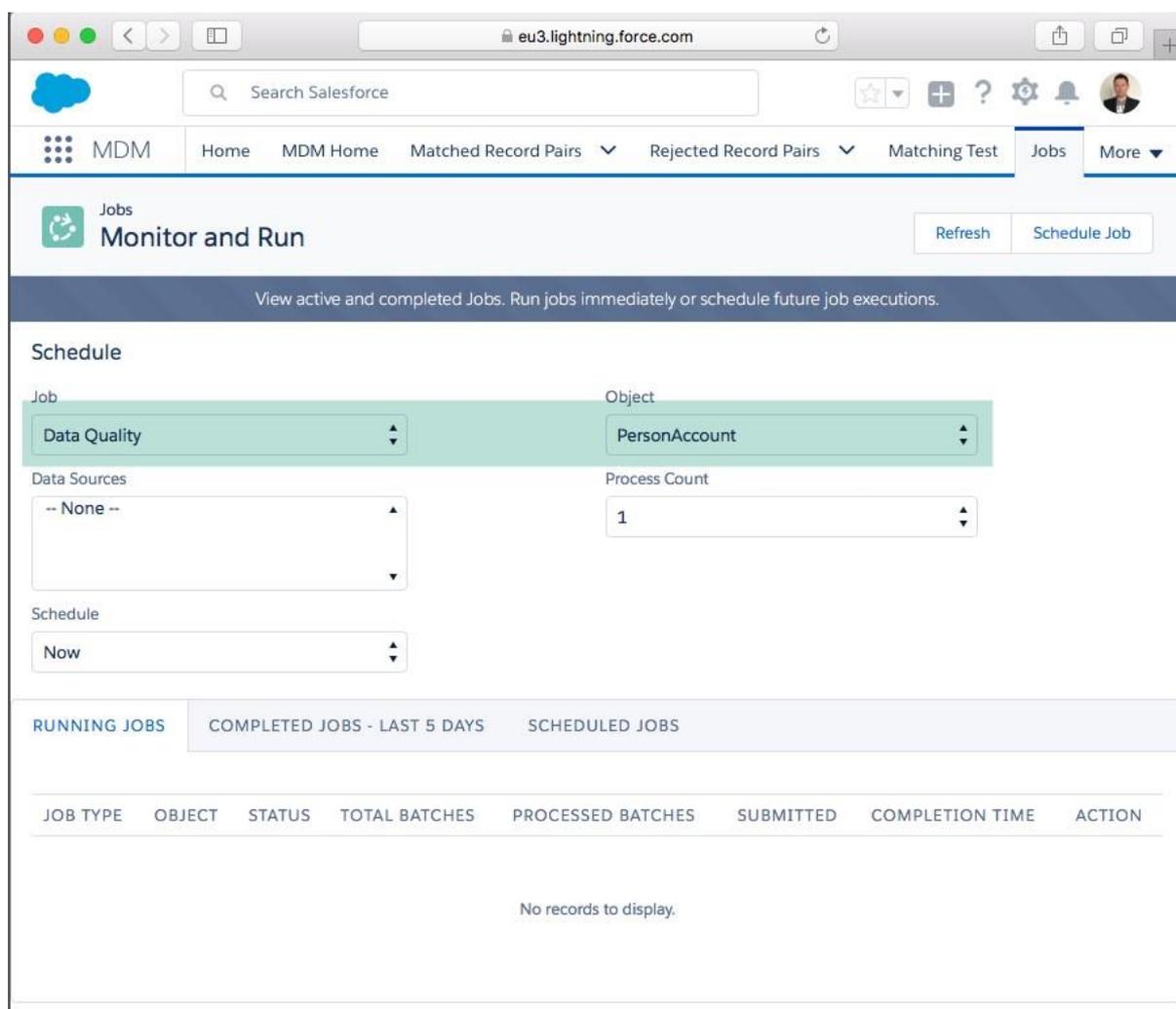
Step 5   View Data Quality Results

The screenshots below show the outputs of Data Quality displayed on an example record. Placement of MDM fields on layouts for administrators, data stewards and end-users is a key implementation decision.

Step 6 Real-time Data Quality – Blocking Actions

The screenshot below shows an example of a blocking action preventing a record modification that decreases the previous quality score. In the example the removal of the email address reduces the quality score as calculated by the Person Account Quality Rule Set.

CLEARMDM

Appendix A – Data Quality Settings Reference

Target Object Data Quality Settings are configured on the Target Object settings page.

Setting	Definition
Is Active?	If set to False the Target Object will be excluded from all Data Quality MDM operations.
Last Quality Check Date Field Name	A datetime field on the Target Object that will be populated with the processing timestamp.

Data Quality Settings are configured on the Data Quality Rule Set settings page.

Setting	Definition
Rule Set Name	A unique name for the Rule Set. Examples; Account Address Quality Lead Score Manage Senior Flag
Order	For each Target Object, Rule Sets are evaluated in a defined order to enable the output of one Rule Set to be referenced as an input to the next Rule Set etc. For example; the standard "Account Address Quality" Rule Set updates the "Account Quality Score" field; this score value is then referenced in the "Account Quality" Rule Set that calculates the overall Quality Score for the record.
Target Object	The Target Object to which the Rule Set is related.
Is Batch Active?	If set to False the Rule Set will be excluded from all batch Data Quality MDM operations.
Is Real-time Active?	If set to False the Rule Set will be excluded from all real-time Data Quality MDM operations.
Data Quality Rules	
Is Active?	If set to False the Rule will be ignored when the Rule Set is evaluated.

CLEARMDM

Rule Type	The template rule type applied to the Source Field; where the field value is in compliance the Score Value is added to the Rule Set score. **Populated**: Field contains a non-blank value. **Minimum Length**: Field length is at least the defined value. **Date Offset**: Field value is within a date range from the current date. **Number Range**: Field value is within a numeric range (e.g. 1-10). **Phone E164 Format**: Field value complies with the E164 phone number format. **Email Format**: Field value conforms to the Salesforce definition of a valid email address. **Exact Match**: Field value is exactly equal to a defined value. **Contains**: Field value contains a defined value.
Source Field	The Target Object field to which the template rule type is applied.
Number Range From	Numeric Range; Starting value of the range.

	Date Offset; Number of days before the current date. Minimum Length; Minimum length value.
Number Range To	Numeric Range; Ending value of the range. Date Offset; Number of days after the current date.
Rule Value	Exact Match and Contains; Defined value to check for.
Score Value	The numeric score value to add or deduct to the evaluated Rule Set score when the Source Field value complies to the defined rule. The score value allows different weightings to be applied to the facts represented by the defined rules.

Reference Rules	
Is Active?	If set to False the Rule will be ignored when the Rule Set is evaluated.
Rule Type	The template rule type applied; where the record is in compliance the Score Value is added to the Rule Set score. **Related Records**: Child records exist that meet the defined filter. **Reference Compliance**: Field value exists in a pre-defined list.
Source Field	**Reference Compliance Rules Only** The Target Object field to which the template rule type is applied.
Reference Custom Setting	**Reference Compliance Rules Only** The Custom Setting from which the list of valid values will be taken from the Name field.
Child Relationship Name	**Related Records Rules Only** The Child Relationship to be checked. For example if the Target Object is Account, then the Child Relationship may be Cases or Opportunities.

Child Object Field Name	**Related Records Rules Only**
	The Child Object Field to which a Filter is applied when checking for the existence of specific records
	For example; for Open Cases the field would be set to [Closed] and the Rule Value would be set to False.
Rule Value	**Related Records Rules Only**
	The Filter Value to applied to the Child Object Field when checking for the existence of specific records.
	For example; for Open Cases the field would be set to [Closed] and the Rule Value would be set to False.
Score Value	The numeric score value to add to the evaluated Rule Set score when the Source Field value complies to the defined rule.
	The score value allows different weightings to be applied to the facts represented by the defined rules.

Data Quality Actions	
Is Active?	If set to False the Rule will be ignored when the Rule Set is evaluated.
Action Type	The template action type applied to defined ranges of the evaluated Rule Set score.
	Field Update: A field update can be applied using the calculated score value, a field-to-field copy or literal value.
	Post to Chatter: A chatter post can be created for notification purposes.
	Examples;
	1) score value is 1 to 5 a quality score field is set to 'Low'.
	2) score value is 5 to 10 a quality score field is set to 'Medium'.
	3) score value is 10 to 20 a quality score field is set to 'High'.
	4) if the score value is above 15 a notification is posted.
Use Score Value	**Field Update Actions Only**
	The Field Update value is set to the evaluated Rule Set score.
Target Field Name	**Field Update Actions Only**
	The Target Object field into which the Field Update value is set.
Source Field Name	**Field Update Actions Only**

	The Target Object field from which the Field Update value will be taken (i.e. field-to-field copy).
Field Update Value	**Field Update Actions Only** The literal value (such as "High") to which the Field Update value will be set.
Score Value From Score Value To	Actions are executed only where the evaluated Rule Set score is within the numeric score range defined for the Action.
Block Actions	
Is Active?	If set to False the Rule will be ignored when the Rule Set is evaluated.
Block Create?	If set to True record creation will be blocked if the evaluated Rule Set score is within the defined range for the action.
Block Update?	If set to True record modifications will be blocked if the evaluated Rule Set score is within the defined range for the action – and the score value is less than the value found in a referenced field which must hold the previous score value. This approach enables record modifications to be blocked only when the score is decreased.
Block Reference Field Name	**Block Update Actions Only** The Target Object field which holds the previous score value.
Score Value From Score Value To	Actions are executed only where the evaluated Rule Set score is within the numeric score range defined for the Action.

Appendix B – Troubleshooting

All clearMDM MDM operations log activity (Start and End times etc.) and errors to the Audit Log Entry object. The MDM application includes an **Audit Log Entry** tab to provide convenient access to this data. Each Audit Log Entry record is time-stamped and related to the parent Batch Job Run record: all MDM operations that run via the Job Method created a Batch Job Run record that records job statistics and status. Monitoring of the Audit Log should be a frequent activity performed by the Administrator or Data Steward responsible for the clearMDM implementation.

Best Practice: Salesforce Reporting Notifications provide a proactive means of reporting on errors generated by clearMDM operations. In this model a standard report is used to return data from the Audit Log Entry object where required conditions are met. Only when records meet the criteria is a report sent to the Administrator or Data Steward responsible.

CLEAR MDM

Data Quality Log Types

Rule Type	Definition
Missing or Inactive Target Object Data Quality settings	This error can occur where the Target Object Data Quality settings are inactive, deleted or the Salesforce User does not have permissions to the object or fields referenced.