

Training – Normalisation Module

Document Version: v1.0

Date: June 2018

Updated: October 2019

Updated: August 2021

This document provides initial training on the Normalisation function within clearMDM .This document does not go into technical detail and is intended to compliment other clearMDM training modules. Existing experience of clearMDM is not required.

Table of Contents

Training – Normalisation Module	1
Normalisation Introduction	2
Step 1 – Standardising the fields	2
Normalisation Rules	3
Step 2 – Setting the rules up in clearMDM	5
Step 3 – Blocking Key	5
New Features	6
Abbreviated Fields	6

Normalisation Introduction

Normalisation is the first function in the clearMDM process. It has two purposes. The first is to standardise the field values, and the second is to produce a blocking key which will be used in the matching function (more about this in the Matching Training Module).

What does 'Standardise the field values' mean? Well, it's probably easier to provide an example to explain this. When record information is captured, depending on who is inputting it and in which system, there could be variations of a field content.

Example:

A Billing address could have a Billing Country text field, which is free format. If this was for the UK, you could have Great Britain, England, Wales etc. By standardising that field so any of these inputs would be classed as 'UK', will help with the matching process.

Also, a company name may have different variations, Barista Coffee, Barista Coffee Ltd, Barista Coffee Limited, or even misspelt, Barrista Coffee. Again, by standardising that field so any variations are recognised, it all helps with the matching process.

So how do you apply this within clearMDM and what's the Blocking Key? Follow these steps and exercises to find out.

Step 1 – Standardising the fields

Decide up front before implementing this within clearMDM, which fields would need standardising. Once this has been decided, which may need input from various Business Members, then decide what the rule should be per field.

This involves understanding what rule type can be used within clearMDM. The table below goes through these and the next step will show where they are selected within clearMDM. For now, concentrate on understanding each one and applying to the fields identified.

Normalisation Rules

Rule Type	Definition
Lookup	This type of rule is used where a custom setting has been setup to reference the input field value and standardise this field. An example for this is the BillingCountry field. The custom setting would have a match list of all possible variations this could take and what this should be standardised to, including the field that should be used to do this.

The following exercise puts this into context.

Exercise:

1. Within your Salesforce sandbox, navigate to setup (top tier level).
2. On the lefthand side, scroll down to Build, Develop and select Custom Settings.
3. clearMDM already comes with a list of pre-determined Custom Settings.
4. Click on Manage 'ISO Country Codes'.
5. Here you will see a short list of countries. You can click on new to create a new one, or select one of the existing ones.
6. For this exercise, click on Edit next to United Kingdom.
7. Within here, you can add into the Match List all possible Country names.
8. The country code will be standardised to GB in the ISO Country Field where any records have the value that is in the match list.

Rule Type	Definition
Remove End	A Custom Setting can be referenced to return a list of values to remove from the input field value.

Exercise:

1. Navigate again to the setup, custom settings area as in the above exercise.
2. Click on company suffixes and manage.
3. In this example you will see a list of values that are to be removed from the input field value. Where a company has Limited, Ltd, Plc after the name, this will be removed so just the name remains for matching.

Rule Type	Definition
Email Domain	This literally removes the domain name from the email address.
URL	Again, this removes the domain name from the URL address.
Format	If format is selected then this will remove whitespace and any special characters (e.g. *{}>) from the target field.
Title	If this rule is selected then it makes sure the text in the targeted field is Title cased, for example The shop = The Shop.
Title Hyphen	This rule will ensure each word in the text for the targeted field is not only Title cased, the words are seperated by hyphens not spaces. The Shop will be The-Shop.
E164	This rule takes the input Phone field and formats the value to the E164 format , e.g. +447730780803. This is determined by a custom setting. The value is displayed in the Normalised Phone field within the record.

Exercise:

1. Navigate to Setup, Develop, Custom Settings are previous.
2. In this example, click on ISO Country Codes, United Kingdom.
3. Click on edit.
4. You will see 3 fields that you can enter the following for:
 - International dialling code for the source country.
 - Local dialling code for the source country.
 - International dialling code for the target country (e.g. +44)

Step 2 – Setting the rules up in clearMDM

Now that the fields and the rules to be applied to these fields have been determined, then this information needs adding into clearMDM.

Exercise:

1. Within your Salesforce Sandbox, navigate to Settings.
2. Click on Target Objects.
3. Select the Target Object that these rules apply to. For this exercise use Person Account.
4. Scroll to the bottom of the page where the Rules section is.
5. Within this section, add the target fields and rules pre-determined in step 1.
6. Ensure 'The is active?' is checked.
7. The match type needs to be set to fuzzy.

Step 3 – Blocking Key

Once the standardised fields have been set, the blocking key also needs to be set. This is also known as the Matching Key and is the first step in providing the initial grouping of records to which matching rules are applied. When a record is normalised, it has the Blocking Key set, based upon characters taken from up to 3 fields. If a record does not have a blocking key, it is not matched at all, so it is very important the blocking key is set correctly.

When setting the blocking key, you will be asked to identify up to 3 fields to take this information from. Ideally this needs to be fields that do not change frequently, i.e. Last Name. Then you will be asked to complete the number of characters to be taken from that field.

Exercise:

1. Within your Salesforce Sandbox, navigate to settings.
2. Select Target Objects.
3. Click edit next to the Target Object, in this case Person Account.
4. Select the relevant input field name from the dropdown.
5. Enter the length of the blocking key for that input field (how many characters do you want to take from the field content).
6. Enter this also for the other two blocking key input fields, or leave as it is if only one field applies.
7. Save the changes.

The normalisation standard rules and the blocking key have now been setup. To run the normalisation job, refer to the Jobs Training Module.

This concludes the Normalisation Training Module.

New Features

Abbreviated Fields

It is now possible to update fields where there has been an abbreviation entered, so an example could be 'Street' abbreviated with 'Str' or 'St'. By setting up a custom setting list of match and replace values and referencing that via a lookup rule, the value of the field will be replaced with Street (in this example) where these abbreviations are used. Within Develop, Custom Settings there is 'Replace Values' where you can add a list of values to be replaced.