

QuickStart Guide 7 – Data Services

Document Version: v1.0

Product Version: v3.13

Date: 13th November 2018

This document provides an overview and Step-by-Step implementation instructions for the clearMDM Data Services 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

QuickStart Guide 7 – Data Services	1
Data Services Introduction	2
Definition.....	2
Key Concepts.....	2
Data Service Methods.....	4
Step 1 – Configure Application Settings	5
Step 2 – Create the Data Service Settings	6
Step 3 – Ensure Target Object Source Records have the Data Services fields displayed	7
Appendix A – Data Services Settings Reference	9

Data Services Introduction

Definition

Data Services within clearMDM provides the ability for record cleanse, verification, or improved quality services via the use of External API connectors.

Fields are added to the Data Service Setting which in turn will be monitored for any changes made within the record source. Any updates to these fields will generate an update request to the external API connector.

This update can be applied one of two ways; transactional or scheduled.

If the Data Service is set to transactional (within Data Service settings), then as soon as the field changes value, the update request will occur. Otherwise, the update is scheduled to run at a certain date and time.

Key Concepts

Concept	Definition
Target Objects	<p>A compatible object that is defined as the target for Matching and Merge 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.</p> <p>A broad range of Standard Objects (including Person Accounts) are supported as both Target Objects and Data Sources. Custom Objects are also supported.</p>
Data Sources	<p>A compatible object that provides data to MDM operations. Each Data Source has a Source Object and a Target Object setting.</p> <p>Internal Data Sources expose data held in the Target Object. Data Source (Account) > Target Object (Account)</p> <p>External Data Sources expose data held in a different object. Data Source (ERP Companies) > Target Object (Account)</p> <p>Partition Data Sources enable a single object to support multiple Data Sources with distinct settings. Partition Data Sources are typically used to group records relating to an external system (e.g. SAP, Sage X3) or to isolate records at different quality grades (e.g. High, Medium and Low).</p> <p>Master Record Data Sources allow Master Records to be exposed to MDM operations via a distinct Data Source with appropriate settings such as elevated merge field priorities. A Master Record Partition data source is implemented as a partition data source that references the record MDM status value.</p>

Master Records	<p>The Master Record is the output of MDM processing and is optionally related to the underlying Source Records (or duplicates) via relationship field or simply through the concatenation of record identifiers into a field on the Master Record.</p> <p>Most typically a Master Record is a de-duplicated Account, Person Account, Contact or Lead record enriched with data from its related Source Records. Where Source Records (or duplicates) are not removed, it is typical to use the Salesforce sharing model to present Salesforce end-users with access to Master Records only – thereby removing the visibility of duplicates.</p> <p>clearMDM implements a custom merge engine that works across object boundaries and can create as well as update Master Records.</p>
Source Records	<p>Source Records are the input to MDM operations and may be related to a Master Record. Source Records can be retained or deleted. Source Records can be considered as the underlying duplicates that are often hidden from view for Salesforce end-users or deleted entirely once processed.</p>
Data Services Settings	<p>Data Services are configured per Target Object on the Data Services settings page. A setting-by-setting definition is provided in Appendix A.</p>
Monitored Fields	<p>Fields which are to be monitored for changes can be added to the Data Service via the settings page. An update request will be generated if any changes are made to these monitored fields in the record source.</p>

Data Service Methods

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

Method	Definition
UI	Within the Data Services Settings there is a field for 'Is Transactional?'. By setting this flag to true, when changes are made on the monitored fields, a Data Service Update request is performed.

Data Services Walkthrough

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.

To complete this step, first navigate to the MDM App, open the Settings tab, tick the two fields Is Active? And Triggers Active? And save.

Application Settings

Save

Application Settings control the global behaviour of the clearMDM application.

Application Settings

Required Information

Is Active?

☒

Triggers Active?

☒

Is Hierarchies Active?

☒

Persist No Matches?

☐

Show Help Captions?

☒

Display System Settings?

☒

Audit Log Level

None

Delete Log Entries?

☐

Delete Audit Log Days

0

Max Jobs Display Days

5

Max Search Results per Data Source

10

Search Term Minimum Character Count

4

Default Fuzzy Match Threshold Percentage

65.00

Manual Field Merge Priority

15

Max Records Per Data Export

5,000

Max Child Record Search Results

200

Max UI Visible MRG Size

200

NamespaceApexPrefix

clearmdm

NamespaceComponentPrefix

clearmdm__

System Settings

Force Deferred Mode

☒

Deferred Dispatcher Mode

☐

Scheduled Job Start Time Delay Seconds

5

Predictive Partitioning Mode

☐

Max MRP For Manual Merge

15

Max Dynamic Hierarchy Parent Records

5,000

Max Blocking Key Groups Merged Per Job

10,000

Max Records Per Group

200

Max Groups Per Iterable Process

10,000

Max Records Per Group From UI

50

Max Records Per Iterable Process

30,000

Max Groups Per Iterable Cycle

15

Max Records Per Iterable Cycle

1,000

Max Records Per Iterable Process Limit

60,000

Use Match Cycle Partitioning?

☒

Max Match Cycles Per Iterable Cycle

50,000

Max Predicates Per Non Iterable Process

200

Match Cycle Processing Split

95

Selective UI Matching Enabled?

☐

Master Record Partition Allowance

50

Save

Step 2 – Create the Data Service Settings

To complete this step, first navigate to the MDM App, open the Settings tab and select the Data Services Tab.

Click on the Add Data Service button, which will display the Data Service Setting page. Within this page, enter the details for the Data Service as explained in [Appendix A](#)

Ensure Monitored Fields have been added to the setting. This can be a single field or many.

Once all the details have been entered, click on Save. This will display the Data Service in the list. Always ensure Is Insert Active? and Is Update Active? Are set to True.

The screenshot shows the 'Data Service Settings' page for a service named 'DSTest'. The page is part of the CLEARMDM application, with a navigation bar at the top including 'Home', 'MDM Home', 'Matched Record Pairs', 'Rejected Record Pairs', 'Matching Test', 'Jobs', 'Batch Job Runs', 'Audit Log Entries', 'Settings', and 'Data Service Update Requests'. The 'Settings' tab is active.

On the left, there is a 'Recent Items' list with user names and IDs. Below it is a 'Recycle Bin' button.

The main content area is titled 'Data Service Settings' and includes a 'Save' button, a 'Delete' button, and a 'Cancel' button. A yellow banner at the top of the settings area states: 'Data Services provide record cleanse, verification or enrichment services via External API Connectors. Hover the mouse over each field to view field-specific inline help.'

The settings are organized into two main sections: 'Data Service Settings' and 'Fields'.

Data Service Settings:

- Name:** DSTest
- Target Object:** Person Account
- Is Insert Active?:** ☒
- Is Transactional?:** ☐
- Is Update Active?:** ☒
- Connector Apex Class Name:** clearmdm.RecordDataSvc
- Max Requests Per Connector Call:** 100
- Filter Field Name:** -- Select --
- Filter Field Value:**
- Is Active For Update Field Name:** Is Active For Update?
- Last Update Date Field Name:** Last Data Services Update
- Field Monitoring Override Field Name:** -- Select --
- Delete Incomplete Requests?:** ☐
- Delete Incomplete Request Days:** 0
- Delete Completed Requests?:** ☒
- Delete Completed Request Days:** 0

Fields:

A yellow banner at the top of the Fields section states: 'Click Add to relate a Target Object Field to this Data Service for Change Monitoring.'

Action	Target Field Name	Is Active?
Del	First Name	<input checked="" type="checkbox"/>
Del	Last Name	<input checked="" type="checkbox"/>
Del	Mobile	<input checked="" type="checkbox"/>

At the bottom of the Fields section is an 'Add' button.

Step 3 – Ensure Target Object Source Records have the Data Services fields displayed

When setting up the Data Service, there is the option to set the 'Is Active for Update' Field Name' and 'Last Update Date Field Name' for the Source Record. These fields will need adding to the Source Record template.

Navigate to the Target Object Records. Select a record and then Edit page layout. Drag the two fields into the MDM section of the template. Ensure the 'Is Active for Update' set to true. Save the changes.

The screenshot shows the 'MDM Data Quality Information' and 'MDM Information' sections of a record. The 'MDM Data Quality Information' section includes fields for 'Quality Score' (1), 'Quality Status' (with a red status icon), 'Address Quality Score', and 'Address Quality Status' (with a red status icon). The 'MDM Information' section includes fields for 'Is MDM Processed?' (checkbox), 'MDM Status', 'Blocking Key' (SMITHE), 'Blocking Key Group' (SMI), 'Last Data Services Update Date', 'Is Active For Update?' (checkbox, checked), 'Is Normalised?' (checkbox), 'Is Blocking Key Complete?' (checkbox), 'Is Active For Matching?' (checkbox), 'Is Active For Reparenting?' (checkbox), 'Is Active For Conversion?' (checkbox), 'Is Conversion Master?' (checkbox), 'Contact Normalised Last Name' (SMITHE), 'Contact Normalised First Name' (CLAIRE), 'Is Master Record?' (checkbox), 'Master Account', 'Last Normalised Date' (22/05/2018 13:45), 'Last Matching Date', 'Last Merged Date', 'Last Conversion Date', 'Fan Ids', and 'ECommerce Customer Ids'. The 'MDM Actions' section includes fields for 'Matching on Save?' (checkbox), 'System Master Field Map', 'Normalise on Save?' (checkbox), and 'Contact System Master Field Map'.

Step 4 – Set the update request to be near real time

There are two methods of scheduling the update request one of these is real time. To do this, within the Data Service settings, select the checkbox 'Is Transactional?' so it is set to true (I.e. ticked).

The screenshot shows the 'Data Service Settings' page for a service named 'DSTest'. The 'Data Service Settings' section includes fields for 'Name' (DSTest), 'Is Insert Active?' (checkbox), 'Is Transactional?' (checkbox, checked), 'Connector Apex Class Name' (clearmdm.RecordDataSe), 'Filter Field Name' (dropdown), 'Is Active For Update Field Name' (dropdown), 'Field Monitoring Override Field Name' (dropdown), 'Delete Incomplete Requests?' (checkbox), 'Delete Completed Requests?' (checkbox), 'Target Object' (Person Account), 'Is Update Active?' (checkbox), 'Max Requests Per Connector Call' (100), 'Filter Field Value' (dropdown), 'Last Update Date Field Name' (Last Data Services Updat), 'Delete Incomplete Request Days' (0), and 'Delete Completed Request Days' (0). The 'Fields' section includes a table with columns 'Action', 'Target Field Name', and 'Is Active?'. The table contains three rows: 'Del' for 'First Name' (Is Active? checked), 'Del' for 'Last Name' (Is Active? checked), and 'Del' for 'Mobile' (Is Active? checked). There is an 'Add' button at the bottom of the table.

Note: To set this as a scheduled job, untick the 'Is Transactional?' Check box.

Step 5 – Make a change to a monitored field

In this example a change will be made to the last name. Navigate to the record, make the change and save. The 'Is Active for Update' checkbox will be ticked indicating an update request has been made. The update request will be displayed in the Data Service Update Requests Tab as new, and as it is set to run almost immediately, will shortly display as complete. At this point , the update request has been sent.

MDM Information	
Is MDM Processed? <input type="checkbox"/>	Is Master Record? <input type="checkbox"/>
MDM Status <input type="checkbox"/>	Master Account <input type="checkbox"/>
Blocking Key <input type="checkbox"/> SMITHE	
Blocking Key Group <input type="checkbox"/> SMJ	Last Normalised Date <input type="checkbox"/> 22/05/2018 13:45
Last Data Services Update Date <input type="checkbox"/>	Last Matching Date <input type="checkbox"/>
	Last Merged Date <input type="checkbox"/>
	Last Conversion Date <input type="checkbox"/>
Is Active For Update? <input checked="" type="checkbox"/>	
Is Normalised? <input type="checkbox"/>	
Is Blocking Key Complete? <input checked="" type="checkbox"/>	Fan Ids
Is Active For Matching? <input type="checkbox"/>	ECommerce Customer Ids
Is Active For Reparenting? <input type="checkbox"/>	
Is Active For Conversion? <input type="checkbox"/>	
Is Conversion Master? <input type="checkbox"/>	
Contact Normalised Last Name <input type="checkbox"/> SMITHE	
Contact Normalised First Name <input type="checkbox"/> CLAIRE	

Appendix A – Data Services Settings Reference

Target Object Data Services Settings are configured on the Data Services Settings page.

Setting	Definition
Data Services Settings	
Is Active?	A checkbox which if set to False the Data Service is not active for processing new records.
Is Update Active?	A checkbox which if set to False the Data Service is not active for processing updated records.
Is Transactional?	A checkbox which if set to True , the Data Service Update Requests are processed in near realtime. If it is set to False , the class clearmdm.Data-ServiceUpdateRequestProcessor must be scheduled via the Apex Scheduler (in Setup)
Target Object	A drop-down containing all Target Objects. Select the one that is associated to this Data Service Setting.
Name	A text field for the unique name of the Data Service.
Connector Apex Class Name	A text field for the Connector Apex Class Name associated with this Data Service Setting.
Max Requests Per Connector Call	A drop-down from 1 – 1000 to select the maximum number of Update Requests per connector call.
Filter Field Name	A drop-down containing a list of fields on the Target Object which is used to identify the subset of records to be processed by this Data Service.
Filter Field Value	A text field to enter the Filter Field Value.
Is Active for Update Field Name	A drop-down boolean field (True or False) containing a list of options that indicates whether an individual Source Record is pending Data Service updates. Is Active for Update? Should be selected.
Last Update Date Field Name	A drop-down containing a list of date fields on the Source Record which is indicates when the last Data Service Update Operation took place. Last Data Services Update Data Date should be selected.
Field Monitoring Override Field Name	A drop-down Boolean field (True False) containing a list of options for a checkbox field on the Target Object that triggers a Data Service Update when the field changes from False to True. If set this setting overrides configured data Service Fields which will be ignored.
Delete Completed Requests	A checkbox which if set to True , the Data Service Update Requests at Completed status are deleted. If this is set to False then the records will be retained indefinitely.
Delete Incomplete Requests	A checkbox which if set to True , the Data Service Update Requests at New or Pending status are deleted. If this is set to False then the records will be retained indefinitely.
Delete Completed Request Days	A drop-down field containing 1 – 30 to select the number of days after the Record Creation Date that Completed status Data Service Update Requests are deleted. A zero value indicates the same day.
Delete Incomplete Request Days	A drop-down field containing 1-30 to select the number of days after the Record Creation Date that New or Pending status Data Service Update Requests are deleted. A zero value indicates the same day.
Fields – click Add to relate a Target Object Field to this Data Service for Change Monitoring.	
Target Field Name	A drop-down list of all the fields in the Target Object Records to select from making one the Target Field Name.
Is Active?	If set to True indicates this field is active for this Data Service Setting.

Data Service Update Request Tab

Setting	Definition
New Data Service Update Request	
Data Service ID	A text field to enter the unique identifier for the Data Service.
Owner	A Non-editable field pre populated with the user creating the Data Service.
Data Service Name	A text field to enter the unique Data Service Name.
Target SObject Type	A text field to enter the name of the Target Object.
Status	A text field to enter the status of the Data Service as either new, Pending or Completed.
Record ID	A text field to enter the record identifier.
Connector Notification Date	A date and time selector field when connectors were notified that the related Record completed Data Service processing.
Record Name	A text field to enter the name of the Record.
Blocking Key Value	A text field to enter the Blocking Key value.