

data8 duplicare™ installation

Table of Contents

Prerequisites	3
Installation	4
Configuration	5
Initial Configuration	5
Duplicate Detection	6
Creating a Dedupe+ Rule	6
Setting up the details of the Dedupe+ Rule	7
Viewing the duplicate results.....	14
Viewing the duplicates with the custom control	14
Viewing the duplicates with Duplicates Detected+	15
Merging.....	15
Creating a Merge+ Rule	15
Field Merge Rules.....	16
Advanced Settings.....	17
Merging Pipeline	18
Technical Details	19
Enable QuickMerge.....	19
Global Search	20
Enabling Rules for Global Search	20
Configuring Mappings	21
Using Global Search	22
Dedupe+ Jobs.....	24
Creating a dedupe+ Job	24
Processing the results	25
Dedupe+ Job Schedules	27
Address Standardization	29
Creating an Address Field	29
Technical Details	31
Telephone Number Standardization.....	32
Creating a Phone Number Field	32
Technical Details	33

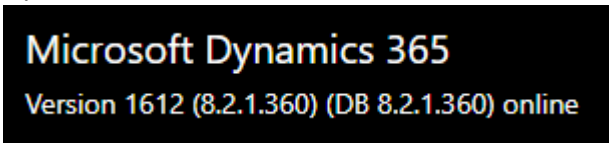
Lookup+.....	34
Creating Rules for Lookup+.....	34
Configuring Mappings.....	34
Configuring your Form.....	35
Using Lookup+.....	36

Prerequisites

Before installing the data8 duplicare™ solution, please review the following list of prerequisites:

1. **Microsoft Dynamics 365**

To confirm your product version, click on the gearwheel icon in the top right corner of the Dynamics screen and click About



Confirm that the first number in brackets starts with at least 8.1

2. **Online or IFD**

Your Dynamics 365 instance must either be hosted by Microsoft (with a domain name ending in dynamics.com), or be configured for Internet Facing Deployment (IFD).

3. **Data8 account**

You must have a valid Data8 account. Please register at <https://www.data-8.co.uk/> if you have not already

4. **Registered CRM connection**

We need to be able to connect to your CRM for the duplicare™ solution to work correctly.

Please enter your connection details at <https://www.data-8.co.uk/Dashboard/CrmConnections>

5. **Assign Privileges**

When registering your connection, you will be given an option of using either a user account or an application user. Please assign the relevant security permissions to whichever user you have opted to use.

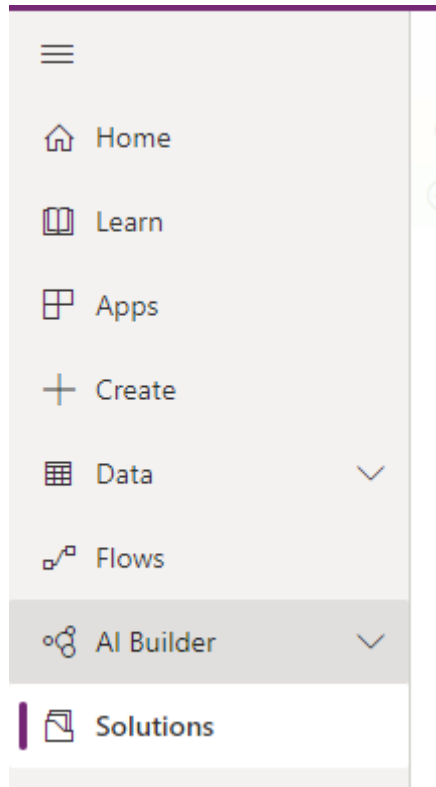
At minimum we will need the duplicare administrator role and the system customizer role to help us configure. On top of that, we will need relevant permissions to perform the actions you want us to perform (i.e. if you are merging accounts, we would need all permissions associated with that - not only in account but in it's child tables etc). Generally, if you're looking at granular permissions it can take some time to get them right which is why we always prefer system administrator. If our systems do fail with missing permissions though, we will get in touch and let you know what we are missing.

Once you have met these prerequisites please contact your Data8 account manager to apply your duplicare™ license to your account.

Installation

Note: If you installed the solution from AppSource then this step will have been completed for you.

You will have been provided with a zip file containing the postcode lookup solution. To install it, open <https://make.powerapps.com>, ensure you're in the right environment in the top right of the page and navigate to "Solutions" on the bottom of the left nav bar.



Click "Import" and select the zip file you have been sent, and then click "Next".

Click "Next" again and the solution will be imported.

Click "Close" and the solution has now been imported.

Configuration

Initial Configuration

All configuration for duplicare can be completed in its own model driven app.

When installed, find your list of “Apps” in the left-hand menu and open the “Duplicare Administration” app.

Enter your Data8 username and password that you use to log into the Data8 website and press Save. This will create the link from your Dataverse/Dynamics 365 to the Data8 systems. Once it says, “Details saved successfully” you can close the pop-up welcome window.

data8 duplicare™

Data8 Username

Data8 Password

Save

After this step is complete, your installation is set up and you are ready to customise the product to satisfy your business needs.

A few other options may appear when you click “Save”.

Data Processing Location:

Use this to control where in the world your data will be processed. By default, all external data processing will be handled in the Data8 data centres, please contact your account manager to enable additional options.

Configure Global Search:

Please refer to the “Global Search” section of this guide for more information on that section.

Toggle Advanced Debugging:

Occasionally issues do arise, and our support team may ask you to turn advanced debugging on, which would help us diagnose any issues.

Duplicates Detected+:

Please refer to the “Duplicates Detected+” section of this guide for more information on that section.

Duplicate Detection

Creating a Dedupe+ Rule

data8 duplicare™ enhances the standard duplicate detection rules with its own Dedupe+ Rules. Navigate to Dedupe+ Rules on the main sitemap and click the “New” button to create a new rule.

Give the rule a name to help you identify it and enter your base & matching record types. This works in the same way as the standard duplicate detection rules – the base type defines the table that will be matched from, and the matching type defines the table that will be found. These should be entered as the internal “logical” names of the table types, e.g. “account” or “contact”.

Real Time Enabled

The “Real Time Enabled” option defines whether this rule should be applied to identify duplicates as a user is entering a row (if set to “Yes”), or only to identify duplicates within your database in bulk (if set to “No”).

Global Search Enabled

The “Global Search Enabled” option allows you to determine if the rule is going to be used anywhere you use a global search. Configuration of global search is mentioned in a later section of this configuration guide.

Records to Search

By default, duplicare will only return potential duplicates that the initiating user is able to see based on your existing security roles. By toggling this, you can allow duplicare to search across all rows in the system. If a duplicate is found in this way and the initiating user should not be able to view the potential duplicate row, they will not be shown the row but will be told a matching row exists. At no time will we overrule your existing security and reveal data that a user is not able to see.

Rule Name	*	Contact Match
Base Record Type	*	contact
Matching Record Type	*	contact
Real Time Enabled	<input checked="" type="checkbox"/>	Yes
Global Search Enabled	<input type="checkbox"/>	No
Records to Search	<input type="checkbox"/>	Records Visible to User

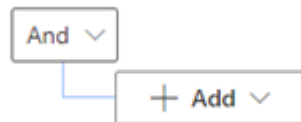
Once you are happy with your selections, save the row. At this point, the solution will make various required customizations to your system to support the duplicate detection system so this can take a little while to complete. Once it has finished saving, the option to define the details of the rules will appear.

Setting up the details of the Dedupe+ Rule

Click the drop-down list to select a rule to add. The following types of rule are available:

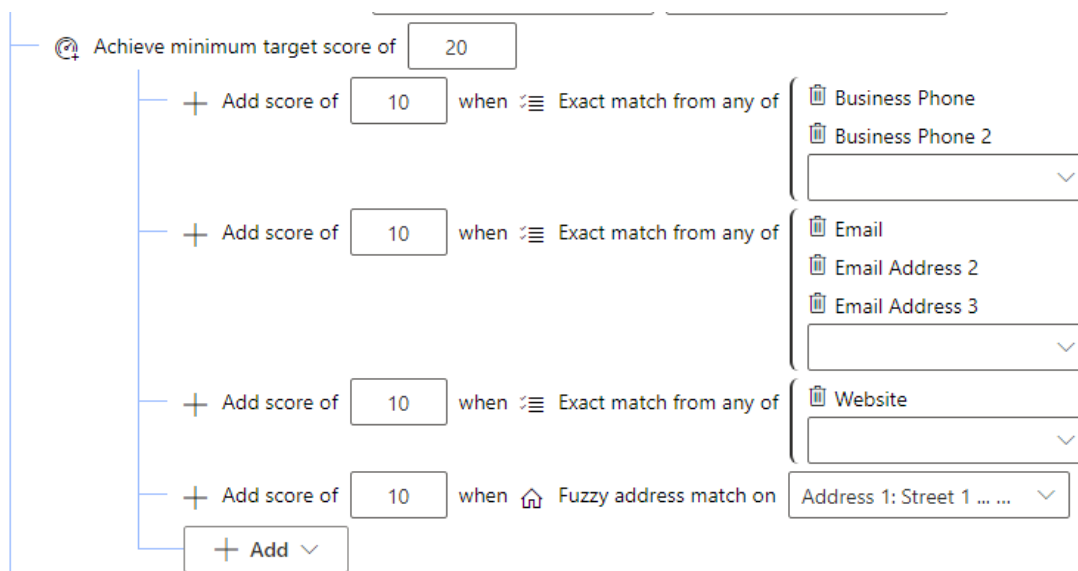
And/Or

Groups other rules together. A default “And” grouping is applied automatically to the rules you add. You can switch an And to an Or by clicking the drop-down arrow to the left.



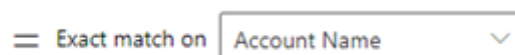
Scoring

Groups other rules together and assign each rule a score, either positive or negative alongside a minimum threshold to make the score pass.

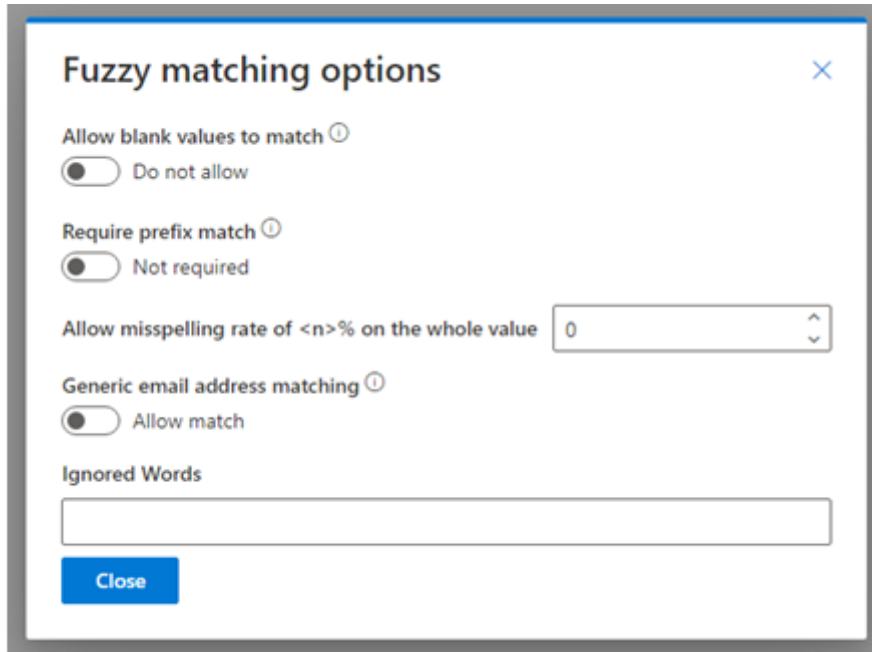


Exact Match

Provides simple matching of exact columns, like the standard duplicate detection rules. Select the column to match.



Additional options are available by clicking the drop-down arrow to the left and selecting “Fuzziness Options” – note the options you see may not all be visible based on the data type you are generating the rule on:



Allow blank values to match:

This will allow a blank value to match another blank value. This does not allow a populated value to match a blank – that is covered by the “Non-Conflicting” rule type below.

Require an exact match on the first x characters:

This will apply the exact match only to the starting characters of the column.

Allow a misspelling rate of...:

You can specify a misspelling rate to allow fuzzy matches in case of mis-typed values. This is expressed as a percentage value which represents the number of characters the two values are different by. If the value on your master row is 10 characters long and you specify a misspelling rate of 20%, this will match rows with values up to 2 characters different.

Ignore Generic Matches on emails:

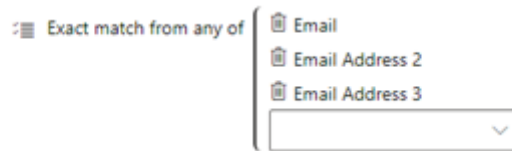
When used on an email column, this will look for the value on the left hand side of the @ symbol and will exclude it if we deem it to be a generic value such as (but not limited to) sales@, marketing@, info@.

Ignored Words:

Each word entered onto it’s own line will be removed entirely from the value. For example if you had “test” in this box then “Data8 test duplicare” and “Data8 duplicare” would classify as an exact match.

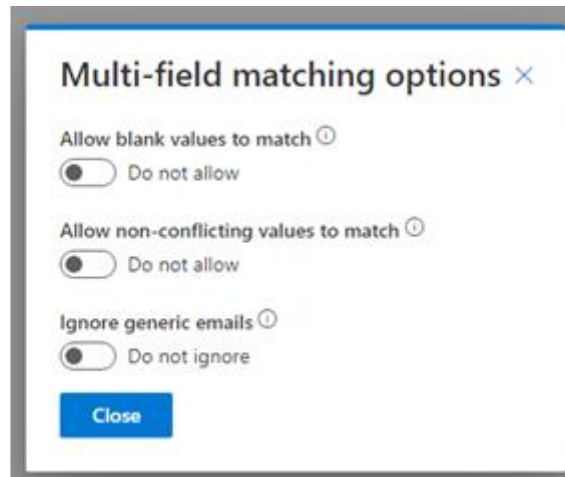
[*Exact Match Across Fields*](#)

Provides matching of exact values across a set of columns. Use this to allow matching of any telephone number or email address for example. In the drop-down list, select each column you want to include in the match in turn to build up the list of columns.



A row will count as a match if any of the columns match, e.g. the Business Phone from one row matches the Home Phone of another.

Additional options are available by clicking the drop-down arrow to the left and selecting “Fuzziness Options”:



Allow all blank values to match:

This will enable rows where all the column values listed are blank to match another row where all the column values listed are blank.

Allow a blank value record to match a record with values:

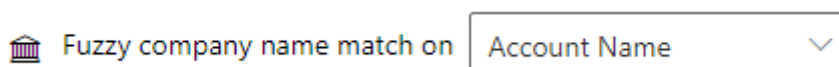
This behaves the same as the later defined “Non Conflict” rule but across multiple columns. A row with populated column values will match a row with all blank column values but will not match a row with different populated column values.

Ignore generic matches:

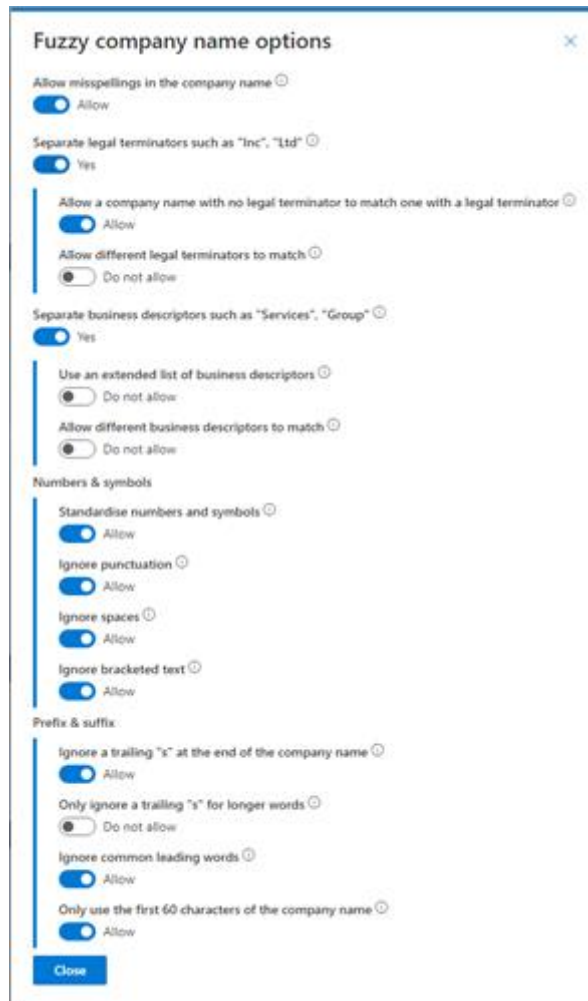
When used on a email columns, this will look for the value on the left hand side of the @ symbol and will exclude it if we deem it to be a generic value such as (but not limited to) sales@, marketing@, info@.

Fuzzy Company Name Match

Provides specialized fuzzy matching for company names, considering the various ways of writing what appear to a human to be the same. For example, this would allow “Data8 Ltd”, “Data 8 Limited”, “Data Eight Group” to match.



Options to control the details of what matches will be allowed can be found by clicking the drop-down arrow to the left and selecting “Fuzziness Options”:



Using this rule will create a standardized company name column automatically which may take a short while to initialize so by amending this rule, you may not see the results reflected immediately.

Fuzzy Address Match

Provides specialized fuzzy matching for addresses, taking into account the various different ways of writing an address that can all relate to the same location.

Before being able to select an address to match, you must first configure the set of columns to be used for the address under the Settings > Address Fields section.



Options to control the details of what matches will be allowed can be found by clicking the drop-down arrow to the left and selecting “Fuzziness Options”:



Allow all blank values to match:

This will enable rows where all the address is blank to match another row where the address is blank.

Allow a blank value record to match a record with values:

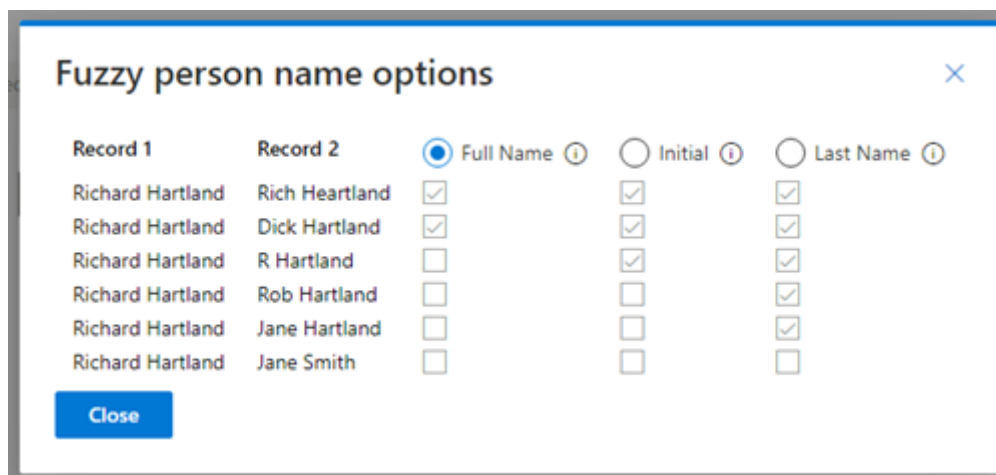
This behaves the same as the later defined “Non Conflict” rule. A row with populated address will match a row with a blank address but will not match a row with a different address.

Fuzzy Person Name Match

Provides specialized fuzzy matching for person names, considering nicknames as well as misspellings. Select the columns that hold the first and last names as well as optionally specifying a title field. The title field isn’t required but will assist in making stronger matches.

Fuzzy person name match on -- Title -- First Name Last Name

Options to control the details of what matches will be allowed can be found by clicking the drop-down arrow to the left and selecting “Fuzziness Options”:



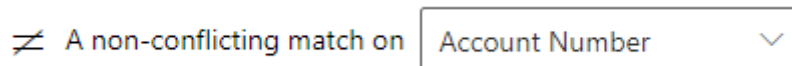
Single Value Match

Restricts the matching to a specific subset of rows, independent of the row being created/updated. Commonly used to match only active rows by selecting the status column but can also be used to apply different rules to different types of rows, e.g. tighter matches on customers vs. prospects.



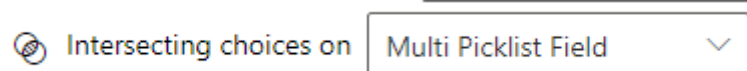
Non-Conflicting

Similar to the Exact Match rule but allows any value to match a blank. Commonly used for columns that provide a link to an external system where merging two rows with different values would cause an error while still allowing manually entered rows to be matched.



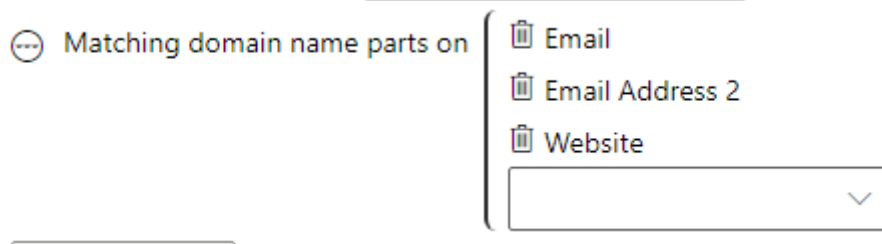
Intersecting Multi-Select Picklist Match

Introduced later into the platform, a multi select picklist allows multiple options to be chosen from an option set. When using this rule, a single common selection between these multi selects will be enough to produce a duplicate match.

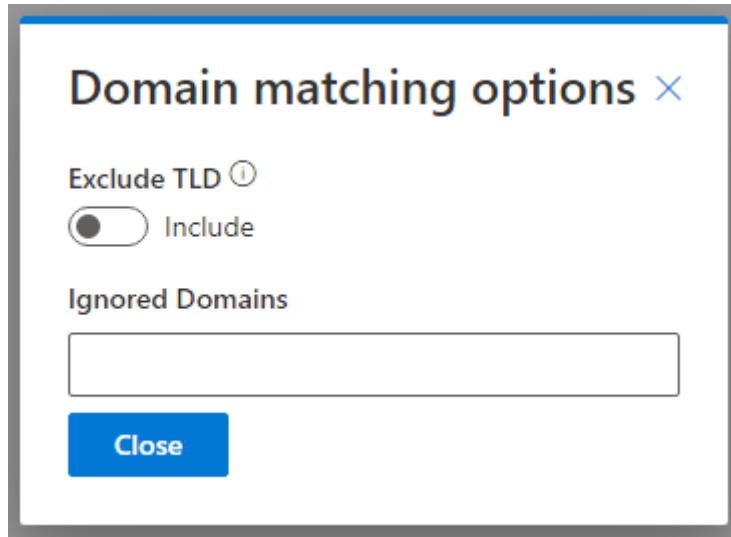


Domain Match Across Fields

Domain match allows multiple columns to be selected and the domain will be extracted out to perform a match. For example, email addresses such as helpdesk@data-8.co.uk would match website urls of <https://www.data-8.co.uk>.



Options to control the details of what matches will be allowed can be found by clicking the drop-down arrow to the left and selecting "Fuzziness Options":



Exclude TLD:

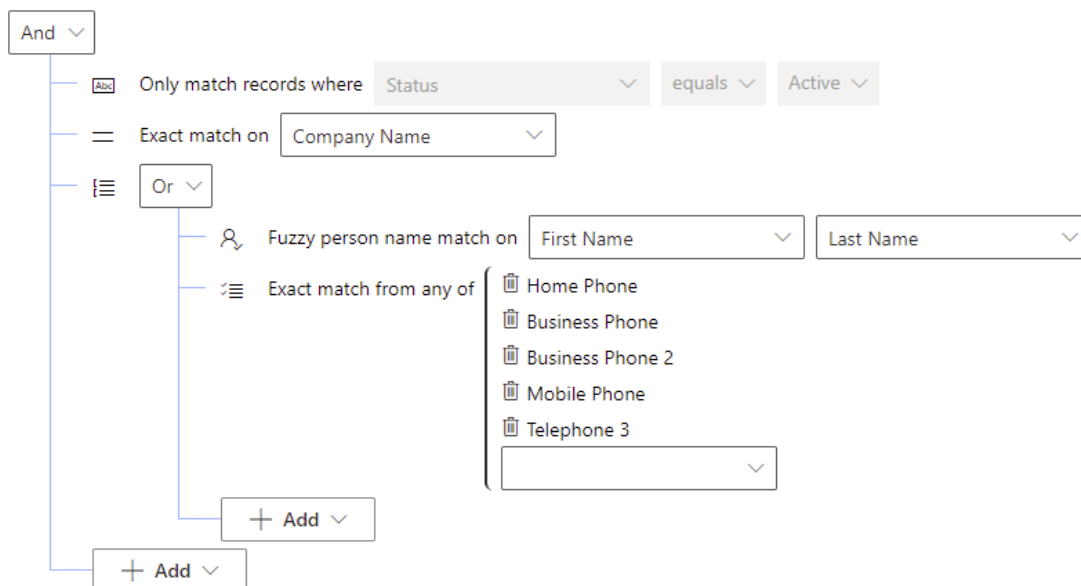
Any match will exclude the top level domain in a match so <https://www.data-8.com> would match mattb@data-8.co.uk.

Excluded Domains:

Any domains that you do not wish to be included in the search – for example you may decide to not do a search on any row with an @gmail.com email address or a website of <https://www.gmail.com>.

Combining The Rules

As an example, we can combine these rules to find matches between active contacts in the same account where there is either a match on the name or telephone numbers as shown below:

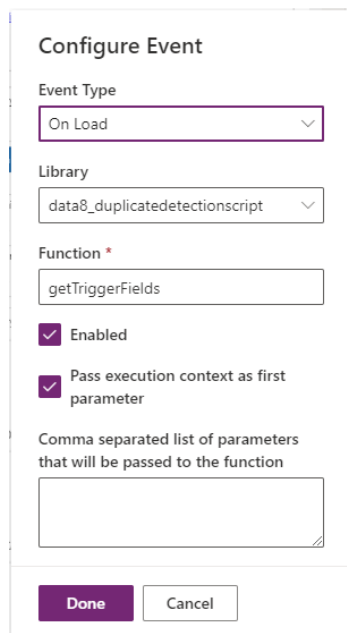


Viewing the duplicate results

When you created the Dedupe+ Rule, a new column “Duplicate Detected Id” (logical name `data8_duplicateddetectedid`) is automatically created on the selected table (both base and matching, if different) and a standard duplicate detection rule is created on it. The new column is automatically removed when the Dedupe+ Rule is deleted, assuming no other rules require it.

To make duplicate detection work you will need to complete these changes manually on the forms you want it to work on. Both Main and Quick Create forms can be customized in this way:

1. Add the “Duplicate Detected Id” column to the form. It can be hidden if required or you can leave it on display and set up the custom control from the next section.
2. Add the `data8_duplicateddetectionsript` script library to the form
3. Add an event handler to the form OnLoad event. Select the `duplicateddetectionsript` library, enter `getTriggerFields` as the function name and enable the option for passing execution context.



Configure Event

Event Type
On Load

Library
data8_duplicateddetectionsript

Function *
getTriggerFields

Enabled

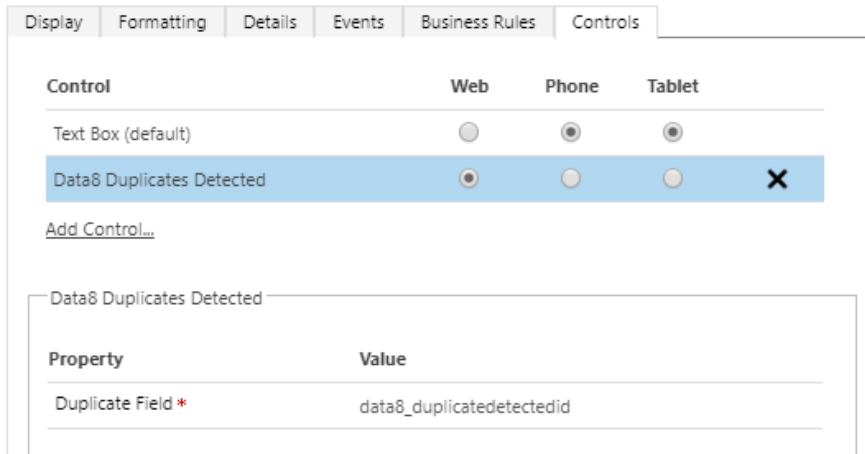
Pass execution context as first parameter

Comma separated list of parameters that will be passed to the function

Done Cancel

Viewing the duplicates with the custom control

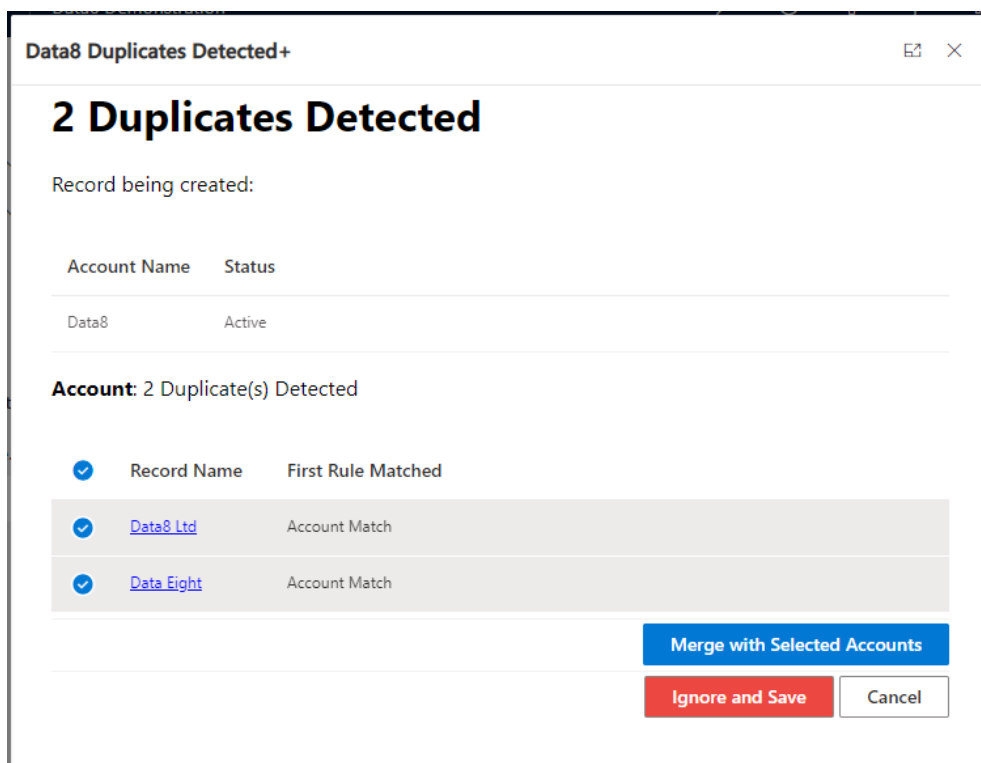
Using the classic form editor, you can assign custom controls to any column – within the duplicate solution is a `Data8 Duplicates Detected` custom control which can be used on the `data8_duplicateddetectedid` column to quickly show duplicates.



Viewing the duplicates with Duplicates Detected+

Introduced in Wave 1 2021 was a “Merge” button on the out of the box duplicates detected window. This clashes with some of the core functionality and to counter this, we recommend you enable “Duplicates Detected+” option from the main authentication page – see earlier section.

By enabling the “Duplicates Detected+” option, the out of the box “Duplicates Detected” will be replaced by a duplicate equivalent which, as well as the columns, will show the name of the dedupe+ rule which was satisfied and found the duplicate.



Merging

Creating a Merge+ Rule

data8 duplicare™ enhances the standard merging functionality with its own Merge+ Rules. Navigate to Merge+ Rules on the main sitemap and click the “New” button to create a new rule.

Enter the table logical name into the “Entity Name” column. Start to enter the name and an auto-complete pop-up with the available options that match the entered characters will show. Select the correct table from the list.

New Merge+ Rule

General Advanced Field Merge Rules

Entity Name * **account**

Save the row. This may take some time as several customizations will be made automatically to the selected table type to enable merging.

Field Merge Rules

The layout that is used is the same as the layout that was applied the last time the chosen table was viewed. Any that are not on the layout will be grouped together in the “Additional Fields” section at the bottom.

If you wish to specify a layout to always be used for the merging layout, that can be done by selecting the option “Use this form for all users”.

Form layout

Contact▼

Use this form layout for merge window

Yes

Each column can have any number of merge rules applied to it. If none is applied, the value of the column is left unchanged from the master row. To apply a Merge Rule to a column click the column and modify the “MERGE RULES” in the right side of the window.

The following merge rule types are available (some are only applicable to specific types of columns):

Use value from master record / Use first value from sorted records

A value (by default, the value from the selected column) is taken from the first row in the list. A different source column can be selected if required.

Use most common value

The set of rows being merged is analysed to identify the most common value from the specified source column. By default, the source column will be the same as the selected target column but can be changed if necessary.

An optional filter can be applied to the set of rows being merged, e.g. to exclude blank values. In this case, only those rows matching this filter are analysed.

If there are two or more different values with the same number of occurrences, you can choose to

continue to the next rule or to pick an arbitrary winner. If continuing to the next rule, only those rows that are part of the potential winning values are passed as candidate rows to the next rule.

Use most common value across a set of fields

This rule only applies to text columns. The set of rows being merged is analysed to identify the most common values across multiple columns (e.g. Email Address 1, Email Address 2 and Email Address 3). Any blank values are automatically excluded. The Priority setting then identifies which of the values to take, e.g. "1" indicates to take the most common value, "2" to take the second most common etc. When different values occur the same number of times, an arbitrary ordering is applied between them.

Concatenate

This rule only applies to text columns. The value is taken from the specified source column from each row being merged and then concatenated together into a single value. Optionally ensures each value is included only once, where multiple rows have the same value, excludes blank values and can use different separators.

Filter & Sort

Takes the set of rows being merged, optionally applies a filter to it to exclude specific rows and then sorts the result. The value is then taken from the specified source column from the row that is first in that sorted list. If two or more rows have the same sort order you can choose whether to pick one arbitrarily or continue to the next rule. If continuing to the next rule, only those rows that are at the top of the sorted list are passed on to the next rule.

Sum

This rule only applies to numeric or money columns. The value is calculated by adding up the value from each row being merged. This is currency aware, so £100 + \$150 with an exchange rate of £1 = \$1.50 will result in £200 or \$300 depending on the currency of the merged row, not £250 or \$250.

Union

This rule only applies to multi select picklist columns. The value will be a list of all values combined.

Intersect

This rule only applies to multi select picklist columns. The value will be a list of values that are selected across all rows.

[Advanced Settings](#)

Merged Status Reason:

This specifies the status reason for the rows that are NOT picked as the master row for tables that do not have "merged" by default. In this box, enter the option set value for the Status Reason column (*statuscode*) for the specified table.

This is optional and if not specified the "default" value for the "Inactive" status reason is used.

Master Selection Rule:

An "order by" clause that uses our "Expression & Sorts" (see [below](#)) and is used to specify the order that the rows are sorted to pick out the Master Record (top record after the sort has been applied).

This is optional and if not specified the "default" master record will be set to the earliest created record.

Show Other Fields:

By default the merge+ screen will display all columns on the form, following by all other columns in an “Others” section at the bottom. If this “Others” section is not required, it can be disabled here.

Merging Pipeline

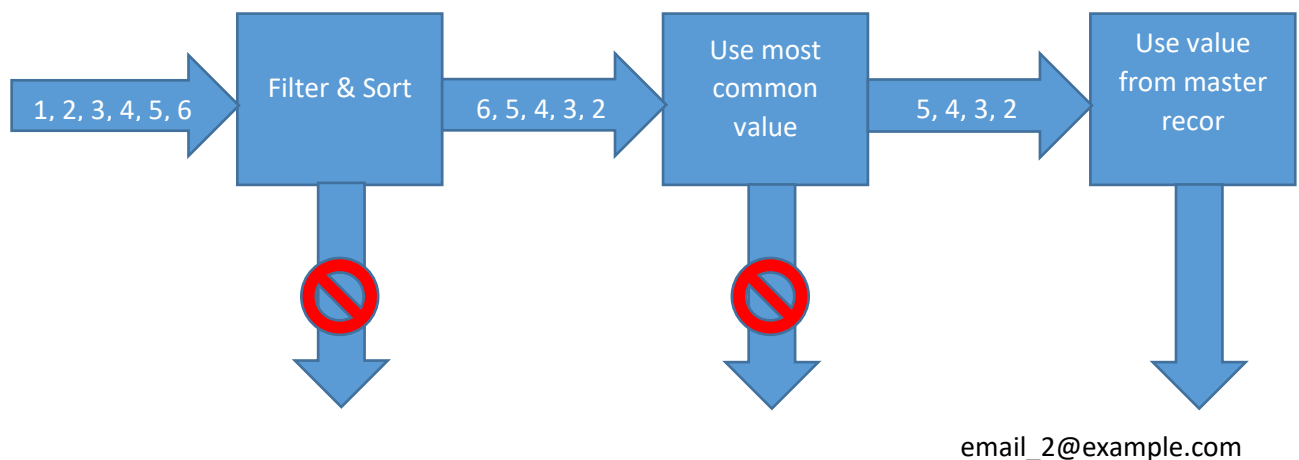
The merge process can be thought of as a pipeline, with the set of rows to be merged being passed from one rule to the next until a rule produces a value. The set of rows may be filtered or reordered in each step, which can affect how the next step operates. The following example attempts to demonstrate how this works:

Example Pipeline

The following rows are selected for merging:

#	Created On	Full Name	Email Address 1	Do Not Email
1	2017-01-01	Joe Bloggs	email_1@example.com	True
2	2017-02-02	Jo Bloggs	email_1@example.com	False
3	2017-03-03	Joe Blogs	email_2@example.com	False
4	2017-04-04	Jo Blogs	email_1@example.com	False
5	2017-05-05	Joe Bloggs	email_2@example.com	False
6	2014-06-06	Jo Blogs	email_3@example.com	False

We want to take the best email address from this set of rows. The master selection rule is set to “createdon” to prefer the oldest row, so the rows will start in the pipeline in the above order.



Filter & Sort is configured with Filter: donotemail = false and Sort: createdon desc. This will filter out row #1. As there are still 5 row left that it cannot distinguish between it does not return a value, but passes those 5 rows on to the next stage in the pipeline. However, the rows are now sorted in the opposite order, with the most recently created row first.

Use most common value will now look for the most common email address. Although email_1@example.com occurs 3 times and email_2@example.com only twice in the original set of rows, row #1 has been filtered out, so there is a draw. Row #6 is eliminated as email_3@example.com only occurs once. Rows #5, #4, #3 and #2 are passed on to the next step (preserving the sort order from the previous step)

Use value from master record will now simply take the value from the first row, i.e. row #5

Expressions and Sorts

Expressions & Sort use a language called **Expressions**. This should be familiar to administrators who are used to SQL or CRM's own calculated columns.

Each expression can reference any of the columns in the rows being merged by their logical names.

Common operators can be applied to the values, such as + - < > = !=. Literal numeric and string values can be used, with strings enclosed in single quotes.

Filters are a single expression that result in a Boolean (true or false) value. This is typically done with a comparison operator such as < > = != is is not, e.g. emailaddress1 is not null. These can also be combined with and and or.

Each sort is composed of an expression and a direction, **ASC** or **DESC**. Multiple sorts can be combined with commas, e.g. emailaddress1 ASC, createdon DESC. The direction can be omitted, in which case it is assumed to be **ASC**.

When you enter a sort or expression and leave the column it is validated, and any validation errors are highlighted in red and must be corrected before saving the changes.

Technical Details

If you create a Merge+ Rule for an table that does not natively support merging (anything other than Accounts, Contacts, Leads and Cases), a new column "Merged Into" (logical name `data8_masterid`) will be created along with a relationship `data8_masterid_entityname` to relate a merged row to its master.

The ribbon for the selected table will be customized to add our own merge-related buttons. This will also hide the default Merge button if the table is one of Accounts, Contacts, Leads or Cases. These ribbon changes will be made in a new unmanaged solution called "DeDupePlus Customizations".

The new column and the ribbon changes are automatically undone when the Merge+ Rule is deactivated or deleted.

Enable QuickMerge

Access to the QuickMerge functionality is controlled by read access to a configuration table – because of this all system administrators by default will have access to QuickMerge but non-system administrators will not.

In order to give Quick Merge permissions to a user, they will need the "data8 duplicare™ Quick Merge" security role. Once applied, QuickMerge will become available on the ribbon.

Global Search

Global search is a feature in which you can run any number of rules at a single time across your system to quickly find existing rows using the data8 duplicate logic.

Enabling Rules for Global Search

Each dedupe+ rule contains a column that says “Global Search Enabled”. If this is set to yes, the rule is included for all users using Global Search.

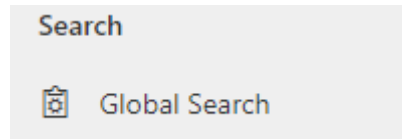


When you change or set this to “Yes” the ribbon of the matching table type will have a button added which looks like this if there isn’t already one there.



If setting the column to “Yes” requires that ribbon change, you will see an “Import Solution” system job appear which is importing that ribbon change. Once that solution has completed importing, perform a “Publish All Customizations” for the button to appear.

We also recommend adding the “data8 Lookup+” ([data8_advancedlookup.index.html](https://data8.com/advancedlookup/index.html)) web resource to your site map as a stand alone web resource for ease of use.



Configuring Mappings

Each column used across all of the global search enabled dedupe+ rules can be mapped for global search. To configure these mappings, navigate back to the configuration page of the solution and click the “Configure Global Search” button which will redirect you to the configuration page.

Search Mappings

data8 duplicare™ Global Search Configuration

<< Back to home

Save Changes

Search Mappings

Display Name	Display Type	Maps To	
<input type="text" value="First Name"/>	<input type="text" value="Text"/>	<input type="text" value="contact"/> <input type="text" value="firstnam"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>			
<input type="text" value="Last Name"/>	<input type="text" value="Text"/>	<input type="text" value="contact"/> <input type="text" value="lastnam"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>			
<input type="text" value="Phone Number"/>	<input type="text" value="Text"/>	<input type="text" value="contact"/> <input type="text" value="telepho"/>	<input type="button" value="Remove"/>
		<input type="text" value="contact"/> <input type="text" value="mobilep"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>			
<input type="button" value="+ Add New Search Mapping"/>			

To configure global search, you can create input columns for the form and map them to columns specified on the dedupe+ rule.

To add a new column entirely, click the “+ Add New Field” button. This will create a new row on your global search. You can also specify the type of the input – at the time of writing, this will not change any behaviour and is more for reference.

Once you have a column row, you can add a new mapping to that column by clicking “+ Add New Mapping”. Two drop down boxes will appear; one will contain a list of possible tables and one a list of the possible columns that can be mapped – note that changing the table can change the columns you have available as this is a dynamic list of columns. Only columns that do not have a current mapping will be listed – if you wish to re-map a column you’ll first have to remove the existing mapping.

Once you have built up your data mapping, click “Save Changes” and it will store that configuration for use throughout the system.

Results Mappings

By default, the global search results window will only include the columns included in the dedupe+ rule that generated a match. There are times when additional columns are required and these can be added in the results mapping section.

Result Columns

Display Name	Entity Name	Attribute Name	
<input type="text" value="Email Address"/>	<input type="text" value="Entity Logical Name"/>	<input type="text" value="emailaddress1"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Results Column"/>			

Start by clicking “+ Add New Results Column” which will generate a new row. Within this row you can enter the following:

Display Name:

This is the header of the column and is just for display purposes only.

Entity Name:

By design, global search can search across multiple tables so displaying the columns in the results can vary based on the type of the result. This column helps specify which specific table that the result is linked to and should contain the logical name of the table. If left blank, the system will try to use it for all the different result types.

Attribute Name:

This is the logical name of the column you wish to display on the global search results. If a table name is specified, it will only be used on that table.

Once you have entered your row of data, click the in line “Save”, before clicking the “Save Changes” at the top to persist all your changes.

Using Global Search

To use global search, either click the “Global Search” button that was added to the ribbon or open the web resource directly. A form should display which is using the configuration you have just made.

data8 Global Search

First Name

Last Name

Phone Number

Input the data you wish to search for and click "Search". This will work through all your global search enabled rules before listing your results.

data8 Global Search Results

1 potential match found



Record Name	Entity Type	Email Address
Data Wizard	Contact	info@data-8.co.uk

Dedupe+ Jobs

A dedupe+ job is a way of taking a dedupe+ rule and applying it across either all rows or a defined subset to locate and solve duplicates in batch.

Creating a dedupe+ Job

Dedupe+ Jobs are stored as rows in the dedupe+ job table and by saving a new row, it will communicate with the Data8 servers and process the job accordingly.

Name	*	---
Owner	*	  Matt Beard
Dedupe Ruleset	*	---
Automerge		<input type="checkbox"/> No
Filtered Master View		---
Filtered Candidate View		---
Store Duplicate Id		<input type="checkbox"/> No
Assignment Strategy		Assign all groups to me

When creating a new dedupe+, the following options are available.

Name:

This is a friendly name for your reference only.

Owner:

This is the owner of the job, again for your reference only.

Dedupe Ruleset:

This is a lookup to the rule set you would like to run against the data.

Automerge:

This is a yes/no option where you can enable the results to be automatically merged. This defaults to "No" and should only be set to "Yes" once you are fully happy with your Dedupe+ and Merge+ rules. If you are running a job with Automerge enabled, we would always recommend creating a back up before as we will not be able to revert our changes.

Run Merge Async: (Online Only)

This option will allow an automerge job to run all it's merges asynchronously as system jobs. Your job complete much quicker and the merges will continue to run behind the scenes. Note – you not be alerted to any failed merges.

Filtered Master View:

Duplicate defines a master as the pot in which duplicates will be looked for. This enables you to reduce that pot so you will only look for duplicates within a subset - for example you only want to find duplicates within your "Active Contacts". The drop-down box for Filtered Master View is populated by both system views and your user views. If left blank, it will use all rows.

Filtered Candidate View:

Duplicare defines a candidate as the pot of rows which we would like to see is duplicates. This enables you to reduce that pot - for example find duplicates of "Leads created on last 24 hours". The drop-down box for Filtered Candidate View is populated by both system views and your user views. . If left blank, it will use all rows.

Store Duplicate ID:

This is optional and defaults to "No". If you select "Yes", the results of the dedupe+ will populate the column you select in the "Store Duplicate ID Override Field" column, enabling to you to perform analysis in a data tool of your choice. Setting this to "Yes" will slow the job down as duplicare will have to apply an update to all affected rows - it will also affect the modifiedon values of those rows.

Store Duplicate ID Override Field:

By default, the duplicate ID will be stored in the "Duplicate Detected ID" column however you can change this option by selecting it here.

Assignment Strategy:

Once the Dedupe+ Job is complete and groups of duplicates have been identified, one or more people will need to work through those groups to merge the rows. If you have many duplicates and need to split this work across multiple people, use this setting to have the groups automatically assigned to the relevant users for review. If you select "Assign groups across a team" then another option will appear.

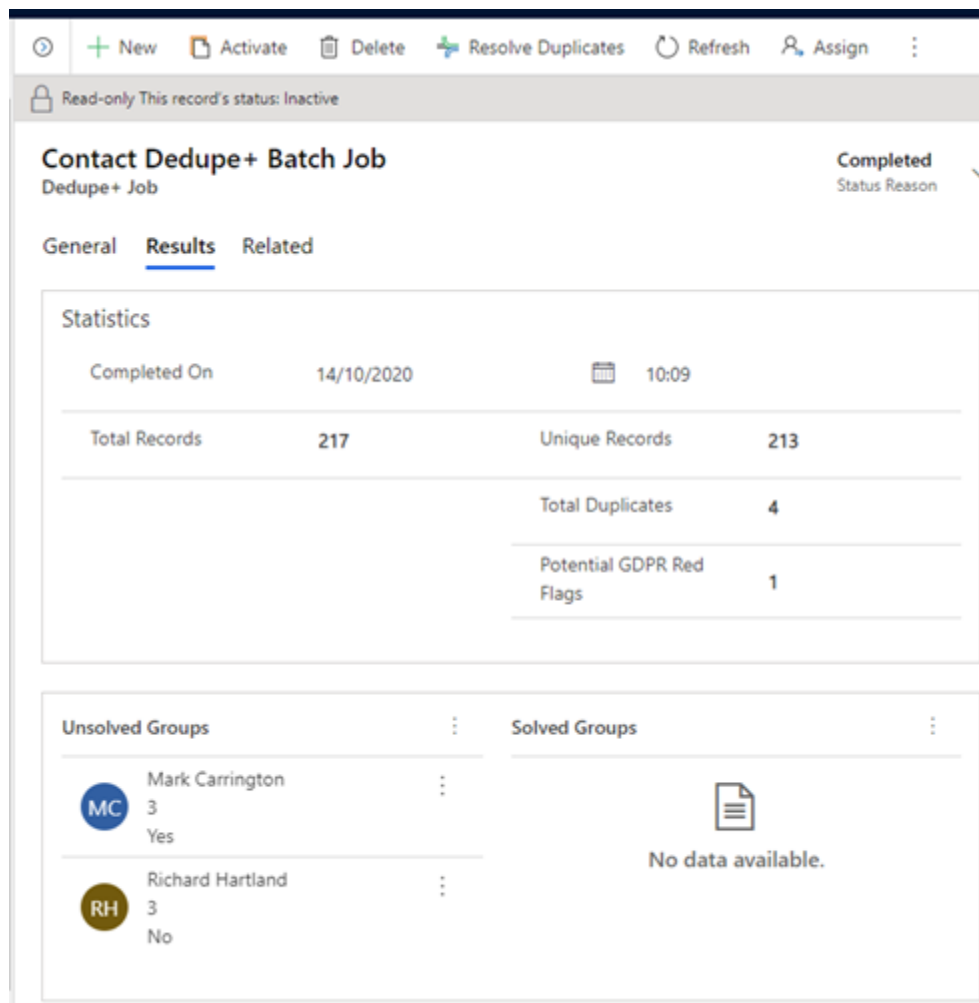
Assignment Team:

This will only appear if you have selected "Assign groups across a team" and it allows you to select the team in charge of handling the duplicate groups.

Once you press save, it will save the row with a default status reason of "Pending". It can take out system up to 5 minutes to pick up the job and you will know we have picked it up as the status will have updated to "Processing".

Processing the results

Once the job is complete, the status will change to "Completed" and if you open the Dedupe+ job row, you will now see a new "Results" tab is visible.



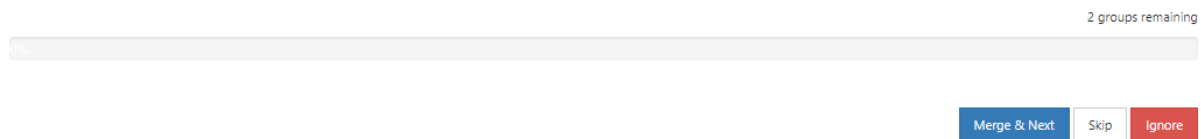
Within this results tab, you have some statistics and some groupings of duplicates. The top line figures are for your reference and will help you understand the level of duplicates you have within your data. Additionally, the "Potential GDPR Red Flags" count you see is to highlight a group of duplicates that contains a conflict of consent i.e. 1 row is "Allow" emails and one is "Do Not Allow". This only refers to the out of the box marketing permission columns and at this time cannot be customised.

As well as the statistics, duplicates will have imported "groups". A group is 2 or more rows we think are duplicated and it's your job to work through those groups and select if it is a duplicate or not and apply that merge. You have 2 options here:

1. Pick a specific group from the bottom grid and by double clicking it, it will open the merging window.
2. Select "Resolve Duplicates" on the top ribbon. This will open the very first group and will enable you to work through each group one by one and make your decisions.

Once you have decided on each groups, you will see that group has moved to "Solved Groups" with a status reason.

When you are merging rows, you will see a slightly different set of buttons at the bottom of the screen which looks like this.



Here you can see how many groups you have remaining, and you can either commit the merge or decide that those rows should not be duplicates.

Merge & Next

This will commit the merge and load the next duplicate group

Skip

This will put this group to the back of the pile for you to make your decision later.

Ignore

This will close this group as a non-duplicate. It will also give you the option to "Would you also like to create all exclusion rules for this group?".

If you click "OK", you will not be told about these potential duplicates moving forward.

[Dedupe+ Job Schedules](#)



You are able to schedule dedupe+ jobs to automatically be generated in order to keep on top of duplicates within your system.





As with all of the duplicare solution, dedupe+ job schedules are stored within a table and can be access through the usual ways, including the duplicare administration app.



To create a new schedule, simply add a new row to this table.

General

Name	*	---
Owner	*	  Matt Beard
Dedupe Ruleset	*	---
Auto Merge	<input checked="" type="checkbox"/>	No
Filtered Master View		---
Filtered Candidate View		---
Store Duplicate Id On Entity	<input checked="" type="checkbox"/>	No
Assignment Strategy		Assign all groups to me

Scheduling		
Frequency	*	Weekly
Schedule Time : Day		Monday
Schedule Time : Hour		00
Schedule Time : Minute		00
 Last Run Time	---	
 Next Run Time	---	

The left-hand section of a dedupe+ job schedule is the same as a standard dedupe+ job and these are the settings that will be copied onto the dedupe+ job when it is automatically generated.

The scheduling section is where you configure the schedule and how often you wish it to run. By populated all these columns and leaving the schedule active you will automatically see the dedupe+ jobs appear.

Please note that all configuration done on this page should be done in the UTC timezone.

Address Standardization

data duplicare™ can standardize the formatting of the addresses on your rows to improve your data quality, make searching for rows and analysing your data more effective, and provide for fuzzy address matching in a Dedupe+ Rule.

Creating an Address Field

An Address Field row groups a set of columns together that represent an address, e.g. Street 1 – 3, City, State/Province and Postal Code. The addresses in those columns are then automatically standardized as rows are created or updated.

Go to Address Fields and click New to get started. Enter the name of the table that holds the address information. This should be the logical name (e.g. “account”, “contact”) of the table – a drop down list will appear as you type to help you out.

Enter the name of the columns in the table that hold the relevant parts of the address. Again, these should be the logical names of the columns and a drop-down list will appear as you type to help. Once you enter the name of the column for Street 1, the others will be populated automatically although you can change them if necessary.

All address fields are required, with the exception of “Company Name”. The optional field you specify for “Company Name” will be used in the fuzzy address match for better matching but will not affect your standardized addresses if you use this feature.

Many addresses will be entered without a country being specified, so use the Default Country Code column to specify the default country that should be assumed where one is not specified on the row. This should be entered as an ISO 2-character code, e.g. “US”, “GB” etc.

The “Standardize” option specifies what should happen with addresses entered these columns. A machine-readable version of the address will be stored in additional columns in the background, but if you set this option to “Yes” a standardized version of the address will also be written back into the same columns it came from.

If you enable “Standardize”, an additional field will be selectable which is the “Standardize Opt-In”. By default with standardization enabled, the address will be re-standardized each time it is edited. If you select a Boolean option in “Standardize Opt-In”, the value of the field selected must be true for standardization to occur – for example you may only want to standardize addresses that are created through Power Automate, so you could set that field during your Flow.

With “Standardize” this turned on your users can enter addresses such as:

Street 1	data8
Street 2	ch2 4ne
Street 3	---
City	---
State/Province	---
ZIP/Postal Code	---

and as soon as they save the row it is transformed into:

Street 1	Data 8
Street 2	Unit 2
Street 3	Venture Point, Stanney Mill Road, Little Stanney
City	CHESTER
State/Province	---
ZIP/Postal Code	CH2 4NE

Once you have created the Address Field row and clicked Save, the status will change to “Initializing”. This may take a while as additional columns are set up automatically on the table to hold the standardized version of the address. As soon as the row is saved, any new or changed addresses will be standardized. At the same time, the existing rows will be extracted, standardized and re-imported into Dataverse/Dynamics 365 to get your historic rows up to the same standard. Once this process is complete, the status will change to “Active”. At this point, it can be used for fuzzy address matching in a Dedupe+ Rule.

Technical Details

When the Address Field row is saved, 15 additional columns will be added to the table. The display name of each one will be “Address 1 (Standardized)” (replacing Address 1 with the display name of the source columns). These are used to hold the various elements of the machine-readable version of the address and should not be referenced directly. They are automatically removed when the Address Field row is deleted.

Telephone Number Standardization

data duplicare™ can standardize the formatting of the telephone numbers on your rows to improve matching of telephone numbers in a Dedupe+ Rule regardless of the punctuation that may be entered around a telephone number, and can also assist with CTI systems that require the telephone number to be entered in a specific format.

Creating a Phone Number Field

A Phone Number Field row identifies the column that holds a telephone number and specifies the type of formatting that should be applied. To get started, go to Phone Number Fields and click New.

Enter the table name that holds the telephone number. This should be logical name of the table, e.g. account, contact, lead. As you type a drop-down list will appear to help you out. Once you've entered the table name, enter the column name as the logical name of the column within that table that holds the telephone number. Again, a drop-down list will be shown as you type.

As most people tend to enter telephone numbers without a country code, enter the Default Country Code as the dialling prefix that should be assumed if one isn't specified, e.g. 1 for North America, 44 for UK.

A new column will be created on the table to store a standardised version of the telephone number. If you want the original telephone number column to be updated with a standardised version, use the following options to control how it should be formatted:

- **Current Country Format Type**
For telephone numbers that are in the same country code as that entered in the "Default Country Code" column, use this formatting.
 - **None**
Leave the telephone number as originally entered by the user.
 - **National**
Change the telephone number into national format, removing any leading international dialling code.
 - **International**
Change the telephone number into international format, adding a + and the default country code if required.
- **Default Format Type**
For telephone numbers that are in any other country code, apply this formatting.
- **Format Punctuation**
If set to "Yes", applies the country-specific punctuation rules that apply to the telephone number. Depending on the country this could involve adding spaces, parentheses, dashes, or other characters to make the telephone number appear in a more natural format to the user.

When you save the Phone Number Field, the row will show in "Initializing" status. At this point, any telephone numbers added or modified will have the relevant formatting applied to them. At the same time, a background job will start extracting all the existing telephone numbers, standardizing them and re-importing them into Dataverse/Dynamics 365. Once this is complete, the status will change to "Active". At this point, the standardized versions will be used for matching in Dedupe+ Rules.

Technical Details

When the Phone Number Field is saved, an additional column will be added to the table. This will have the same display name as the original column but with “(Standardized)” appended to it. Regardless of the formatting options specified, this will always store the telephone number in full international format without any punctuation. This is used by the Dedupe+ Rules and can also be used for CTI systems. This column will be removed when the Phone Number Field is deleted.

Lookup+

Lookup+ is a feature designed to make finding the correct row for a lookup even easier by leveraging the full power of the deduplication logic on data8 duplicare.

Lookup+ allows you to create a bespoke deduplication rule for each lookup, enabling you to quickly find the row you want. If you do not want to use a bespoke deduplication rule but would still like to use Lookup+, it will default to the global search behaviour.

Creating Rules for Lookup+

Lookup+ does not require anything specific on the dedupe+ rule which enables you to make as many rules as you want. The rule does not need to be real time enabled, nor does it need to be global search enabled so you can leave those options off if you wish.

This page will use a similar example throughout - setting the parent account of an account. In this example, we have a column on account that states "Company" or "Site" and the parent account of a "Site" account can only be a "Company" account. This would be difficult to enforce on a standard lookup column.

An example dedupe+ rule has been made like so:

Rule Name	* Account - parentaccountid
Base Record Type	* account
Matching Record Type	* account
Real Time Enabled	<input type="checkbox"/> No
Global Search Enabled	<input type="checkbox"/> No
Records to Search	<input type="checkbox"/> Records Visible to User

Show columns as Display names

Only match active records On

And Only match records where Status equals

Fuzzy company name match on

Only match records where Account Level equals

Translated to plain English, this rule set will find any rows that match the inputted account name, whilst ensuring only active "Company" level rows are returned.

Configuring Mappings

Once you have made the rule you wish to use, you need to make the Lookup+ Configuration - this can be found in the "Duplicare Administration" model driven app included in your solution.

Create a new row and set the columns as required:

- Source Entity Logical Name: This is the logical name of the table you are setting a lookup+ for.
- Source Attribute Logical Name: This is the logical name of the lookup column in which you are setting lookup+ for.
- Dedupe+ Rule: This is a lookup to the dedupe+ rule you have made. If none are selected, the search will default to the global search format.

Once you have set those values, save the form which will refresh the configuration form at the bottom. This is configured almost the same as global search - the only difference is how you save the configuration:

data8 Lookup+ Search Form Configuration

Display Name	Display Type	Maps To		
<input type="text" value="First Name"/>	<input type="text" value="Text"/>	<input type="text" value="contact"/>	<input type="text" value="firstnam"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>				
<input type="text" value="Last Name"/>	<input type="text" value="Text"/>	<input type="text" value="contact"/>	<input type="text" value="lastnam"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>				
<input type="text" value="Company Name"/>	<input type="text" value="Text"/>	<input type="text" value="account"/>	<input type="text" value="name"/>	<input type="button" value="Remove"/>
<input type="button" value="+ Add New Mapping"/>				
<input type="button" value="+ Add New Field"/>				

To add a new column entirely, click the “+ Add New Field” button. This will create a new row on your lookup+ form. You can also specify the type of the input – at the time of writing, this will not change any behaviour and is more for reference.

Once you have a column row, you can add a new mapping to that column by clicking “+ Add New Mapping”. Two drop down boxes will appear; one will contain a list of possible tables that have dedupe+ rules configured and one a list of the possible columns that can be mapped – note that changing the table can change the columns you have available as this is a dynamic list of columns. Only columns that do not have a current mapping will be listed – if you wish to re-map a column, you’ll first have to remove the existing mapping.

Lookup+ will not contain a "Save Changes" button and instead, you are required to save the lookup+ row itself.

Configuring your Form

To set up for your form for Lookup+, you simply need to add a Web Resource in the relevant place - ideally directly underneath the lookup column.

Edit your form, add the "data8_openadvancedlookup.html" web resource onto your form. You can give the web resource any name/label you wish and we recommend settings it to be 1 row high and not displaying a border.

On the general tab, set the properties as following - replacing "parentaccountid" with the lookup column that you are configuring.



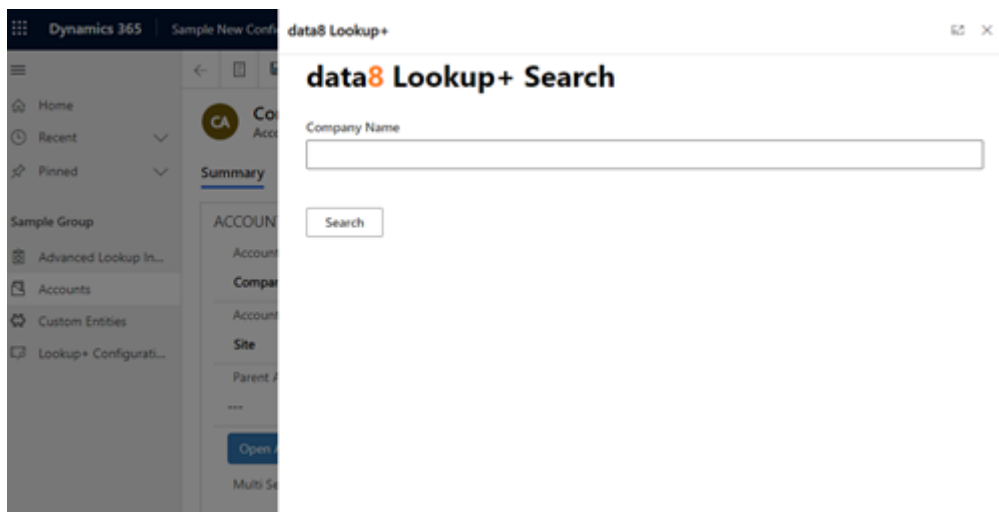
Save, publish and close your form editor.

Using Lookup+

Open an existing row to test out your lookup+ set up. To expand upon the previous example, some sample data has been included as follows:

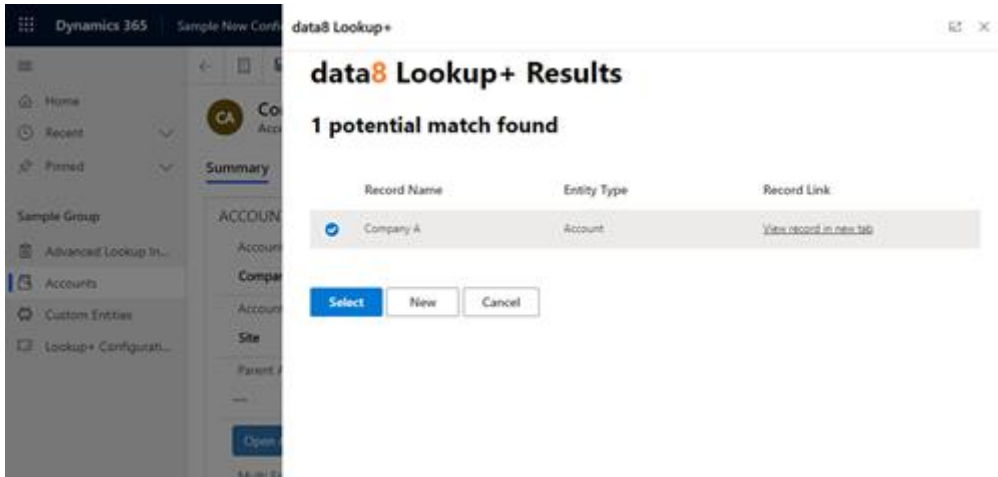
1. "Company" row called "Company A"
2. "Site" row called "Company A (London)"
3. "Site" row called "Company A (Liverpool)"

We would like to check this works as expected by trying to set the parent account of "Company A (London)" and the only suitable candidate for this is "Company A".



Clicking the web resource should pop up the lookup+ search on the right-hand side of the screen. Input your search (in this case, "Company A") and press "Search".

Allow the search to finish to see your results.



Here you can see we have the expected results - we have correctly stopped the "Company A (Liverpool)" row being listed and we are only seeing correct "Company" level rows. If we select the row and press "Select", the column is populated and the search window closes.

