

easyLOOKUP_V2 Setup Guide

Code-Free Fuzzy Database Lookup, Field Correction and Auto-Fill, and Document/Batch Auto-Validation.

Contents

1	OVERVIEW	3
2	FEATURES	4
2.1	MULTIPLE SEARCH CRITERIA	4
2.2	MULTIPLE RUN TIMINGS	4
2.3	FUZZY DATABASE SEARCHES.....	4
2.4	AUTOMATIC DOCUMENT VALIDATION	5
2.5	DYNAMIC BATCH ROUTING - MODULE SKIPPING	5
3	INSTALLATION.....	6
3.1	NEW INSTALLATION	6
3.2	UPGRADING FROM EARLIER VERSIONS.....	6
4	CONFIGURATION	7
4.1	ADDING EASYLOOKUP_V2 TO A BATCH CLASS	7
4.2	MINIMUM REQUIREMENTS	7
4.3	SETTINGS.....	8
4.3.1	Document Class List	8
4.3.2	Batch Level Options	9
4.3.3	Database Settings	10
4.3.4	Lookup Settings	11
4.3.5	Lookup Options	13
4.3.6	No Match Behaviour.....	14
4.3.7	Multiple Match Behaviour	14
4.3.8	Auto Fill Fields	15
4.3.9	Update Settings	15
4.3.10	Remove Settings.....	16
5	GENERAL CONSIDERATIONS	16
5.1	SKIP IF CANNOT LOAD	16
5.2	CONFIDENCE LEVELS AND MULTIPLE FUZZY CRITERIA.....	17
5.3	SQL LOGIC.....	18
5.4	RUNTIME PROCESSING	18
5.5	RUNNING PROCESSES MULTIPLE TIMES	19
5.6	SETTING DOCUMENTS VALID.....	19
5.7	APPLYING EASYLOOKUP_V2 SETTINGS.....	19
5.8	DATA TYPES	19
6	CONTACT.....	20

1 Overview

This guide is intended to show how to install and configure the **easyLOOKUP_V2** workflow agent for Kofax Capture.

It assumes a working knowledge of Kofax Capture batch class setup and an understanding of the standard Kofax workflow process.

easyLOOKUP_V2 is supported for Kofax Capture version 9, 10, and 11.

2 Features

The key features and benefits of **easyLOOKUP_V2** are detailed below.

2.1 Multiple Search Criteria

With standard Kofax Capture database validation a single exact search can be defined:

Ie DatabaseValue=IndexFieldValue

With **easyLOOKUP_V2** it is possible to define up to three search criteria which can be linked with AND/OR to make complex database searches:

```
Ie DatabaseValue1=IndexFieldValue1
    AND
    DatabaseValue2=IndexFieldValue2
    OR
    DatabaseValue3=IndexFieldValue3
```

2.2 Multiple Run Timings

easyLOOKUP_V2 can be run at the close of any Kofax workflow module (apart from Release).

It is possible to run the lookups immediately after Recognition and use the raw OCR/ICR results to search in a database.

It is possible to run the lookups multiple times. For example run after Recognition to see if the OCR results were good enough to get a match. If not, run again after Validation with the user corrected fields.

2.3 Fuzzy Database Searches

With **easyLOOKUP_V2** it is possible to effectively correct OCR/ICR against a database.

A number of Fuzzy search criteria with pre-set confidence levels are available. These allow misread OCR/ICR data to still possibly find a match in a database.

Used after Recognition, a Fuzzy lookup can use the raw OCR/ICR data and check against a database, and populate the field with the correct value from the database.

2.4 Automatic Document Validation

When a match is found in a lookup, it is possible to mark the document valid. This means that there will be no requirement to manually validate the document in Validation.

This can save a large amount of manual effort and cost.

The documents are, however, still available to be checked if required.

2.5 Dynamic Batch Routing - Module Skipping

Where all documents have been marked valid in a batch, it is possible to route the whole batch past Validation dynamically.

3 Installation

3.1 New Installation

To install **easyLOOKUP_V2** follow the steps below:

-Unzip the **easyLOOKUP.zip** file and copy **easyLOOKUP_Setup.dll**, **easyLOOKUP_Runtime.dll** and **easyLOOKUP_V2.aex** to the Kofax Capture Install Path ...\\Bin folder.

NB. For Kofax Capture v10.1 and later, use the files in the Net4 folder.

-In the Administration module, select Tools > Workflow Agent Manager and Add

-Browse to the **easyLOOKUP_V2.aex** file and select it

-Select **easyLOOKUP_V2** and Install

easyLOOKUP_V2 should now be installed and ready for use.

3.2 Upgrading from Earlier Versions

Follow the instruction in 3.1 for a new installation. Once **easyLOOKUP_V2** is installed and registered with Kofax Capture you need to replace the older version on any batch classes which use **easyLOOKUP** processing.

To do this, go to Batch Class Properties > Workflow Agents and add **easyLOOKUP_V2** and then remove earlier versions of **easyLOOKUP**. All previous settings will be retained.

There is an additional setting in this version to distinguish between Access/Excel/Text lookups and other types. By default this will be unchecked. If you are using any of these types of data source, then check that option and Update Settings.

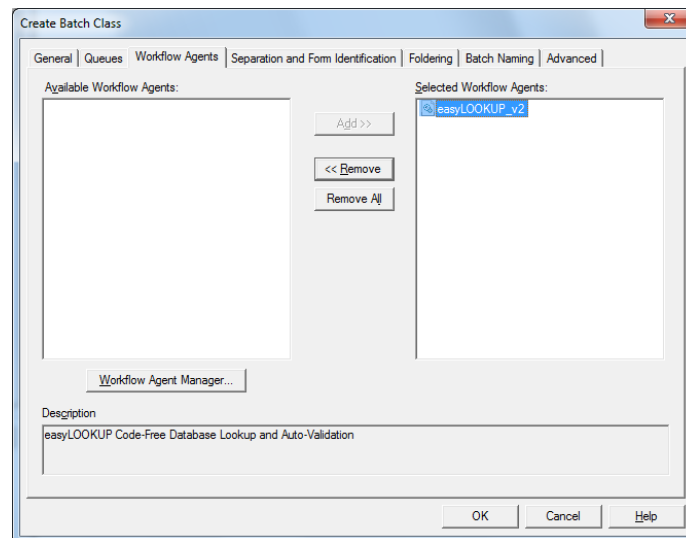
You will need to re-publish any batch classes which have been updated.

You cannot use both versions at the same time.

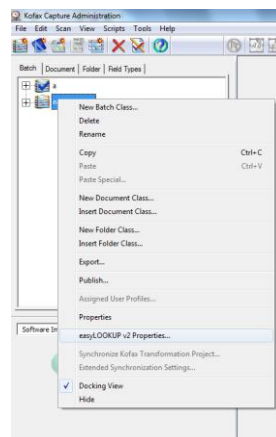
4 Configuration

4.1 Adding easyLOOKUP_V2 to a Batch Class

To add **easyLOOKUP_V2** to a Batch Class, select the Workflow Agents tab from the Batch Class properties, and add **easyLOOKUP_V2** to the Selected Workflow Agents list.



Once **easyLOOKUP_V2** has been added to a Batch Class, right-clicking on the Batch Class will show the **easyLOOKUP_V2 Setup** menu item.

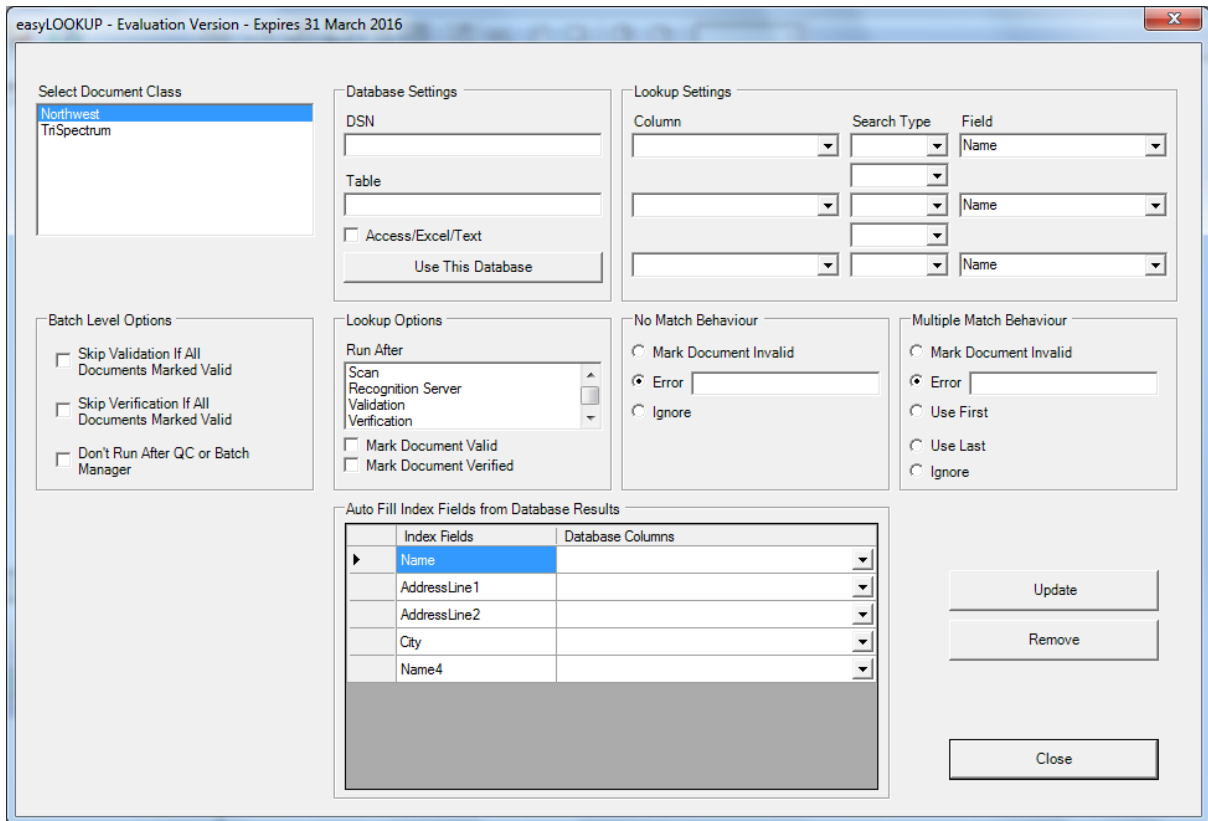


4.2 Minimum Requirements

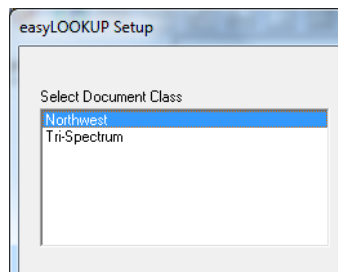
In order to use **easyLOOKUP_V2** a Batch Class must have at least one Document Class defined, with at least one Index Field.

4.3 Settings

The section below details how to configure **easyLOOKUP_V2** processing for a batch class.



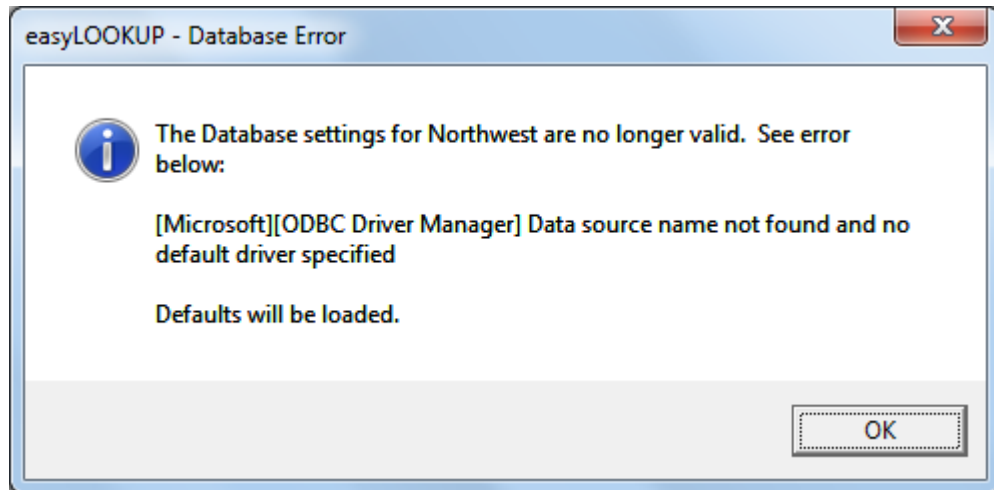
4.3.1 Document Class List



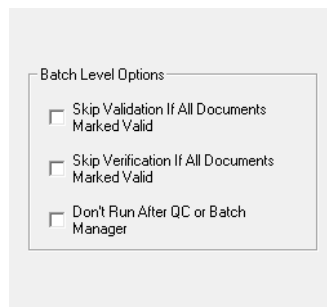
Clicking on a Document Class in the list will load any current settings defined for that Document Class. If no settings are currently defined then default values will be loaded ready for **easyLOOKUP_V2** settings to be configured.

When opening **easyLOOKUP_V2** settings, the first Document Class with any defined settings will be highlighted, and settings loaded.

If the DSN/Table used for settings is no longer available an error will be displayed. Defaults will be loaded for the Document Class.



4.3.2 Batch Level Options



4.3.2.1 Skip Validation if All Documents Marked Valid

If all documents have been marked valid by **easyLOOKUP_V2**, then it is possible to skip the Validation module altogether for a batch. Select this option to use this feature.

4.3.2.2 Skip Verification if All Documents Marked Valid

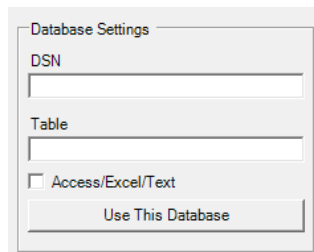
If all documents have been marked valid by **easyLOOKUP_V2**, then it is possible to skip the Verification module altogether for a batch. Select this option to use this feature.

4.3.2.3 Don't Run after QC or Batch Manager

By default a Kofax workflow agent will be called at the close of each module. With **easyLOOKUP_V2** it is possible to restrict that to a certain set of modules by using the Run After list – see section 4.3.5.

However, when manually changing workflow from within Batch Manager, or after encountering an error and going into Quality Control, it may be desirable to skip the **easyLOOKUP_V2** processing.

4.3.3 Database Settings



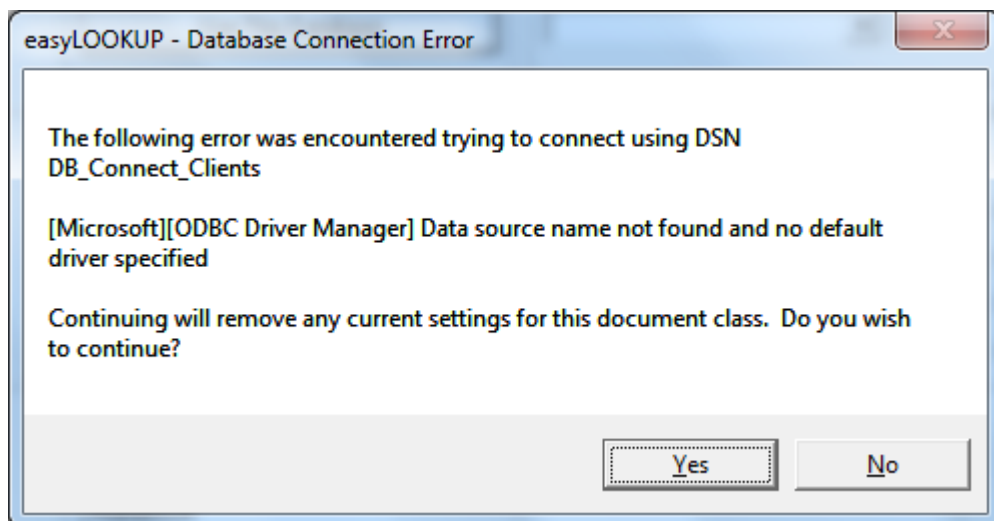
The screenshot shows a dialog box titled "Database Settings". It has two text input fields: "DSN" and "Table". Below these fields is a checkbox labeled "Access/Excel/Text". At the bottom of the dialog is a button labeled "Use This Database".

DSN and Table

Enter the DSN to be used for any lookups on this Document Class, and the Table to be used within the selected data source.

Once entered, click **Use This Database** to check the connection and retrieve the Database Columns to allow lookup configuration.

If there is a problem with the DSN/Table then an error will be displayed.



If the DSN and Table are valid then the column names from the selected source will be retrieved and you can configure the lookup.

Access/Excel/Text

Check this option if the data source you are using is either Access database, Excel spreadsheet, or a plain text file.

4.3.4 Lookup Settings

The 'Lookup Settings' dialog box contains three rows for defining search criteria. Each row has three columns: 'Column', 'Search Type', and 'Field'. All three columns in each row are currently empty dropdown menus.

Up to three search criteria can be defined to create a complex database lookup.

4.3.4.4 Column

The 'Lookup Settings' dialog box shows the 'Column' dropdown menu open, displaying a list of available columns: ID, Surname, FullName, Address, and Phone. The 'Search Type' and 'Field' dropdowns remain empty.

Select a column from the data source to add to the search criteria.

4.3.4.5 Search Type

The 'Lookup Settings' dialog box shows the 'Search Type' dropdown menu open, displaying a list of search types: =, <>, BEGINS, Fuzzy 70, Fuzzy 75, Fuzzy 80, Fuzzy 85, and Fuzzy 90. The 'Column' dropdown is set to 'Surname' and the 'Field' dropdown is empty.

Select a type of search. Available types are:

= An exact match is required against index field value (case insensitive).
 `abc`='abc' True
 `ABC`='abc' True
 `abc`='xyz' False

<> Not equal to. Not matching index field value is required.
 `abc`<>`abc` False
 `abc`<>`xyz` True

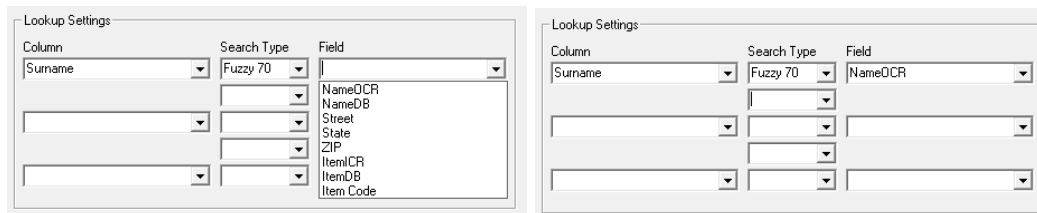
BEGINS The search begins with index field value.
 `abc` BEGINS `ab` True
 `abc` BEGINS `c` False

Fuzzy 70-95 Fuzzy matches with pre-set confidence levels. Only available with Fuzzy version of **easyLOOKUP_V2**

These searches compare a value in the database with the field value and calculate a confidence value based on how close the two values are. If the confidence value is high enough, the value is True, otherwise False. See section 4.2 for detail on how confidence is calculated.

NB. The Fuzzy search requires evaluation of all the possible records in a data source against the Fuzzy matching algorithm. Where using large data lists, it is advisable to use extra search criteria where possible to reduce the processing load.

4.3.4.6 Field



Select which Index Field is to be used in this search criteria. In the above example we are testing the NameOCR Index Field against the Surname column from the data source, and will be testing for a 70% Confident Fuzzy match.

4.3.4.7 Multiple Search Criteria

Column	Search Type	Field
Surname	Fuzzy 70	NameOCR
AND		
Address	Fuzzy 80	Street
AND		
ID	Fuzzy 90	Item Code

By using the AND/OR options, it is possible to build a more complex lookup using up to three database columns and up to three index fields.

See section 5.3 for more details on evaluation logic where AND and OR criteria are mixed in the same lookup.

4.3.5 Lookup Options

Select a module or modules to run **easyLOOKUP_V2** after. Workflow agents are called by the selected module as that module is closed.

In most cases it is advisable to run **easyLOOKUP_V2** after only one module – normally Recognition – however, it can be run after any module other than Release.

Select whether to mark a document valid or verified if a lookup match is found. When a document is marked valid/verified it will not need manual processing in Validation/Verification, but will still be available to be checked by a user if necessary.

If this option is selected in conjunction with the Batch Level option to skip Validation/Verification where all documents are marked valid, then the whole batch can be routed to skip Validation/Verification altogether.

NB. If running **easyLOOKUP_V2** after an interactive module – Scan, QC, Validation, Verification – if the data source is large than can be a delay in closing the batch. For searches against very large data sets it is advisable to run after Recognition, or limit the size of the data set with a more complex search.

4.3.6 No Match Behaviour

Choose an action to take when a lookup returns no results.

Mark Document Invalid	Document will require Validation/Verification if these modules are in the workflow
Error	Reject the document and send batch to QC. The document Note in QC will show the error text defined here
Ignore	No action is taken, document moves to next workflow module

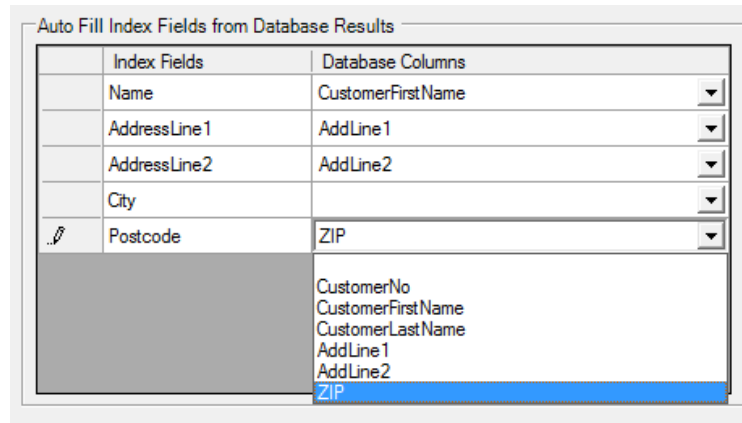
4.3.7 Multiple Match Behaviour

Choose an action to take when a lookup returns multiple results.

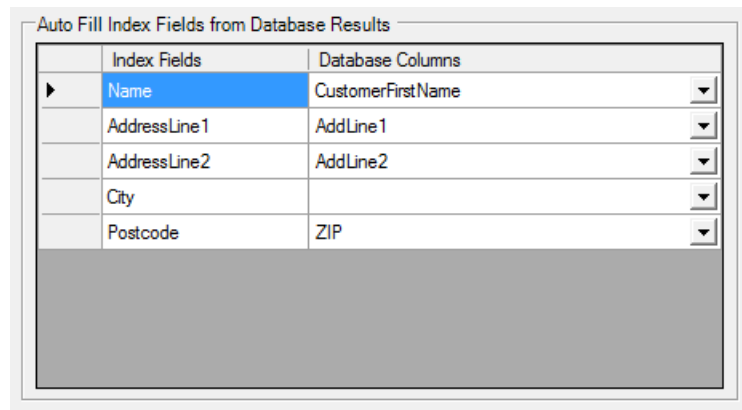
Mark Document Invalid	Document will require Validation/Verification if these modules are in the workflow
Error	Reject the document and send batch to QC. The document Note in QC will show the error text defined here
Use First	The first matching result is used
Use Last	The last matching result is used
Ignore	No action is taken, document moves to next workflow module

NB. When using Fuzzy search criteria, if there are multiple matches the one with the highest confidence will be used. Where there are multiple matches with the same confidence, the settings above will be used to determine action.

4.3.8 Auto Fill Fields



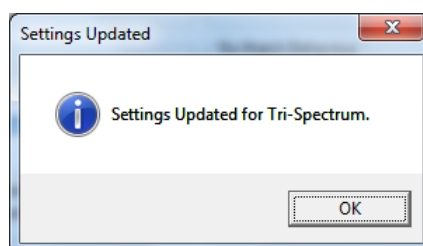
The grid will display a list of the Index Fields for the current Document Class. Clicking in the grid next to a field will allow a selection of the available columns from the database to be used to populate the field after a successful lookup.



Set as many auto fill fields as are required. You do not have to map all Index Fields or all Database Columns.

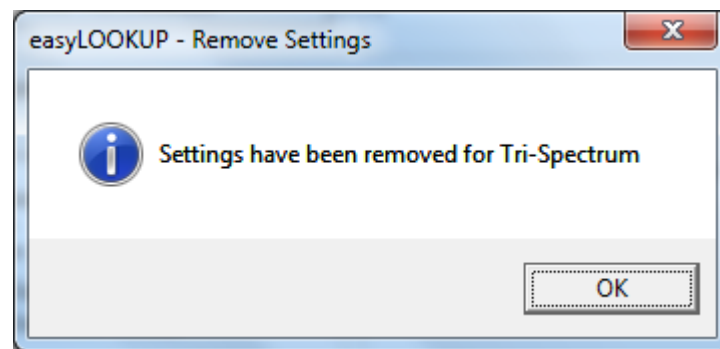
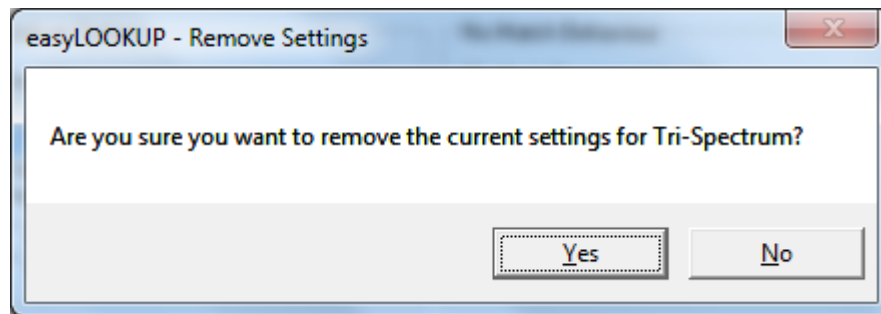
4.3.9 Update Settings

Once settings have been configured for a Document Class, click **Update** to store them.



4.3.10 Remove Settings

To remove settings for a Document Class, click Remove. You will be asked to confirm the removal.



5 General Considerations

5.1 Skip If Cannot Load

An option exists in Kofax workflow agents to skip if not loaded. This option can be set to TRUE or FALSE.

If set to TRUE, if a workstation processes a batch but does not have **easyLOOKUP_V2** installed, it will just ignore the settings.

If set to FALSE, if a workstation processes a batch but does not have **easyLOOKUP_V2** installed, it will cause an error and the batch will be sent to Quality Control.

By default **easyLOOKUP_V2** is set to TRUE. If it is required to set this to FALSE, then the `easyLOOKUP_V2.aex` file needs to be edited so that the highlighted line reads FALSE instead of TRUE. **easyLOOKUP_V2** then has to be re-registered via Workflow Agent Manager in the Administration module. **No other changes must be made to `easyLOOKUP_V2.aex`**

```
[Workflow Agents]
easyLOOKUP_v2
```

```
[easyLOOKUP_v2]
WorkflowAgentID=easyLOOKUP_v2
WorkflowAgentProgID=easyLOOKUP.Runtime
WorkflowAgentFile=easyLOOKUP_Runtime.dll
Description=easyLOOKUP Code-Free Database Lookup and Auto-Validation
Version=2.0
SupportsNonImageFiles=True
SetupProgram=Workflow Agent Setup
WorkflowAgentSkipIfCantLoad=True
```

```
[Workflow Agent Setup]
OCXFile=easyLOOKUP_Setup.dll
ProgID=easyLOOKUP.Setup
Visible=0
MinSizeX=300
MinSizeY=150
BatchClassMenus=easyLOOKUP_V2 Menu
```

```
[easyLOOKUP_V2 Menu]
MenuText=easyLOOKUP v2 Properties...
```

5.2 Confidence Levels and Multiple Fuzzy Criteria

Where Fuzzy lookup functionality is used with a specified confidence level, the confidence level is calculated using a string difference algorithm. This algorithm calculates the number of changes required to match 2 strings. This returns a value known as the string difference. String difference can be considered as the number of additions/deletions/changes required to make string1 the same as string2.

Examples:

1. m4nche5ter and Manchester string difference is 2 (4-a, 5-s)
2. 4nche5ter and Manchester string difference is 3 (add m, 4-a, 5-s)

To calculate the confidence level, the following formula is used:

$$\frac{\text{Length of String (from database)} - \text{String Difference}}{\text{Length of String (from database)}}$$

This would give confidences of 80% and 70% for examples 1 and 2 above.

It is important to note that the greater the difference in length between two strings, the lower the confidence level will be.

Where a lookup uses more than one Fuzzy search, the confidences are averaged to give an overall lookup confidence. Each individual confidence test must pass first.

For example a lookup using a Fuzzy 70 AND Fuzzy 80 would be evaluated as follows:

- If either of the individual Fuzzy criteria is not met, the lookup returns nothing.
- If both Fuzzy criteria are met, then the average of the two confidences is used as the overall confidence for this lookup.

Where multiple matches are found using Fuzzy searches, the one with the highest overall confidence will be used. Where there are multiple matches with the same overall confidence then Multiple Match Behaviour options will be used to determine the action to be taken.

5.3 SQL Logic

When using multiple search criteria in a lookup it is important to understand how the SQL generated will be processed when the lookup is run against a database.

No parenthesis is used in the SQL statement when generated by **easyLOOKUP_V2**.

SQL evaluates search expressions according to set rules of operator precedence. AND is evaluated before OR and similar expressions are evaluated left to right. Some examples of how statements will be evaluated when mixing AND and OR are given below.

Criteria1 AND Criteria2 OR Criteria3
(Criteria1 AND Criteria2) OR Criteria3

Criteria1 OR Criteria2 AND Criteria3
Criteria1 OR (Criteria2 AND Criteria3)

5.4 Runtime Processing

easyLOOKUP_V2 runs at the close of each module in the Kofax workflow, but will only process lookups at the close of modules in the Run After list defined in the setup – see section 4.3.5.

All stations processing Batch Classes where **easyLOOKUP_V2** is used, will need to have **easyLOOKUP_V2** installed unless the *Skip If Cannot Load* option has been enabled. See section 5.1.

Any stations running the modules where **easyLOOKUP_V2** lookups will be processed will need to have the required DSN's and relevant access permissions to any defined data sources.

5.5 Running Processes Multiple Times

It is possible that other workflow agents, or manual actions can force a batch into the same module more than once.

When this happens, **easyLOOKUP_V2** may process the same lookup more than once.

5.6 Setting Documents Valid

When setting documents valid, they will not be automatically opened in Validation and/or Verification.

This is usually the desired behaviour, but it is important to check that no other processes are set to run in Validation or Verification, as these may not run for the documents which have been marked valid.

5.7 Applying **easyLOOKUP_V2** settings

As with other changes in Kofax Capture, a batch class must be published after any changes are made to lookup settings.

5.8 Data Types

Index Fields in Kofax are stored as strings. As the lookups are user-configurable the data type of database columns are unknown. **easyLOOKUP_V2** will attempt to convert the value in the searched database columns in order to perform the match against the Index Field, however in some cases NULL values in a searched field may cause errors. In these cases it is recommended that a view or second table be created to ensure that no NULL values appear in the searched columns.

Ensure sufficient testing before applying **easyLOOKUP_V2** settings to a live system.

6 Contact

For support and assistance please contact mail@davidcrewe.com