Kofax Transformation Modules

Release Notes

Version: 6.3.1

Date: 2019-08-29

KOFAX

© 2006-2019 Kofax, 15211 Laguna Canyon Road, Irvine, California 92618, U.S.A. All right reserved. Portions © 2002-2006 Kofax Development GmbH. Portions © 1997-2006 Kofax U.K. Ltd. All Rights Reserved. Use is subject to license terms.

Third-party software is copyrighted and licensed from Kofax's suppliers.

This product is protected by U.S. Patent No. 5,159,667.

THIS SOFTWARE CONTAINS CONFIDENTIAL INFORMATION AND TRADE SECRETS OF KOFAX USE, DISCLOSURE OR REPRODUCTION IS PROHIBITED WITHOUT THE PRIOR EXPRESS WRITTEN PERMISSION OF KOFAX

Kofax, the Kofax logo, Kofax Transformation Modules, Ascent Xtrata Pro, INDICIUS, Xtrata, Ascent Capture, Kofax Capture, VirtualReScan, the "VRS VirtualReScan" logo, and VRS are trademarks or registered trademarks of Kofax or its affiliates in the U.S. and other countries. All other trademarks are the trademarks or registered trademarks of their respective owners.

U.S. Government Rights Commercial software. Government users are subject to the Kofax standard license agreement and applicable provisions of the FAR and its supplements.

You agree that you do not intend to and will not, directly or indirectly, export or transmit the Software or related documentation and technical data to any country to which such export or transmission is restricted by any applicable U.S. regulation or statute, without the prior written consent, if required, of the Bureau of Export Administration of the U.S. Department of Commerce, or such other governmental entity as may have jurisdiction over such export or transmission. You represent and warrant that you are not located in, under the control of, or a national or resident of any such country.

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

Table of Contents

Chapter 1: About this release....5
Version information....5
New features....5
Natural language processing....5
Changes in behavior....6
Thin Client Server branding....6
Chapter 2: Resolved issues....7
Issues resolved in tools and rich client user modules....7
Reclassify via script....7
Spaces in project names....7
Error deleting protected AZL samples....7
Protected icon missing....7
AZL sample image info....7
Kofax Transformation Modules fix packs....7
Kofax Transformation Modules 6.3.0 fix pack 2....8
Kofax Transformation Modules 6.3.0 fix pack 1....8
Kofax Transformation Modules 6.2.1 fix pack 13....9
Kofax Transformation Modules 6.2.1 fix pack 12....10
Kofax Transformation Modules 6.2.1 fix pack 11....10
Kofax Transformation Modules 6.2.1 fix pack 10....11
Kofax Transformation Modules 6.2.1 fix pack 9....11
Kofax Transformation Modules 6.2.1 fix pack 8....11
Kofax Transformation Modules 6.2.1 fix pack 7....12
Kofax Transformation Modules - Thin Client Server 6.3.0 fix pack 1....12
Kofax Transformation Modules - Thin Client Server 6.2.1 fix pack 12....12
Chapter 3: Known issues....13
A2iA Document Reader Performance Degradation....13
Broken forwards compatibility....13
Clear Thin Client browser history....13
KTM Statistics export connector not installed....13
Named Entity Locator Error....14
Remove Service Pack 6.3.1 on Windows Server OS....14
Root element error when documents copied via drag & drop....15
Chapter 4: Additional Documentation....16

Help updates for Project Builder 16
Updated: Class Details 16
New: Named Entity Locator Properties window 17
New: Sentiment Locator Properties window 19
Updated: Validation sequence 19
New: Manage natural language processing 20
Updated: Database Locator 25
New: Named Entity Locator 25
New: Sentiment Locator 25
Help updates for XDoc Browser 26
New: Nlp 26
Updates for the Installation Guide 26
New: Install Natural language processing engine 27

Chapter 1

About this release

This set of release notes contains important information not included in other Kofax Transformation Modules documentation. Please read these release notes carefully before you install, upgrade, or use this product.

Information about supported operating systems and other requirements is available on the Kofax Support website at www.kofax.com.

Version information

The Kofax Transformation Modules 6.3.1 release has the following build number, which appears in the About Kofax window:

6.3.1.0.0.914.

New features

The following new features are available in Kofax Transformation Modules 6.3.1.

Natural language processing

The Natural language processing engine is now available to extract named entities and the mood or sentiment of a document. This engine is installed separately, and supports several languages. You can use the Named Entity Locator to extract named entities and the Sentiment Locator to determine the sentiment of a document. You can also use a script alone to further process the extraction results, or a script in conjunction with either the Named Entity Locator or the Sentiment Locator to further process their results.

Named Entity Locator

This locator is used to assign extracted entities to fields.

You can choose to extract the entity name only by using a simple field, or you can use a table field to extract not only the entity name, but also the entity confidence, entity type, and the entity sentiment.

Once the entities are extracted, it is then necessary to customize a script to interpret the results for your needs.

Sentiment Locator

This locator extracts text sentiments from a document. This means that the Sentiment Locator is able to determine the overall mood or sentiment of a document based on the words or phrases found on that document.

Once the sentiment is extracted, it is necessary to customize a script to interpret the results for your needs.

Custom entities

You can define your own custom entities. This means that when a specific entity is located on a document, the custom entity file is referenced to determine the extracted result.

Custom entities supports several languages.

Changes in behavior

The following change has been made to the behavior of Kofax Transformation Modules 6.3.1.

Thin Client Server branding

The overall color of the Thin Client Server has been changed from blue to grey.

Chapter 2

Resolved issues

The following are the resolved issues in Kofax Transformation Modules 6.3.1, including previous Fix Packs.

Issues resolved in tools and rich client user modules

The following issues were resolved in the tools or the rich client user modules. These issues are not included in any of the fix packs listed below.

Reclassify via script

When a document is reclassified via script and the result is a valid document, the next document is selected automatically. (1268689)

Spaces in project names

It is now possible to use spaces in project names when creating a project in Project Builder. (1282302)

Error deleting protected AZL samples

It is now possible to delete protected samples from an Advanced Zone Locator without error. (1273832)

Protected icon missing

When a document in the selection classification benchmark set is protected, the protection icon is now displayed correctly. (1269680)

AZL sample image info

It is now possible to successfully create info in AZL sample images without error. (1269588)

Kofax Transformation Modules fix packs

The following fix packs have been released for Kofax Transformation Modules and the Thin Client Server since Kofax Transformation Modules 6.3.0 was released.

Kofax Transformation Modules 6.3.0 fix pack 2

The following issues have been resolved in Kofax Transformation Modules 6.3.0 fix pack 2 (6.3.0.2).

Immediate search breaks in database lookup

When the Immediate search option is enabled, database searches no longer have issues. (1282317)

Restart Service translation truncated

Due to a larger button size, translations of the Restart Services button in Service Configuration is no longer truncated. (1278262)

Incorrect values displayed at batch close

The correct values are now displayed when a batch is closed in the Validation module. (1278126, 1278126)

Merging performance issues in Document Review

There is no longer a performance issue when documents are merged in Document Review. (1261039)

Folder button missing

Users can now successfully navigate between documents when they keyboard shortcuts. (1256889)

Kofax Transformation Modules 6.3.0 fix pack 1

The following issues have been resolved in Kofax Transformation Modules 6.3.0 fix pack 1 (6.3.0.1).

Manual classification not canceled by script

The Classification combo box is now properly updated when an action is canceled via script. (1272474)

Cannot resolve conflicts in extraction set

Conflicts in the Extraction Training Set are now resolved without any issues. (1269877)

Field highlighting still visible for protected document

Highlighting for a protected document is no longer visible in the Document Viewer. (1267356)

Splitting protected documents

It is no longer possible to split a protected document. (1267336)

Protected documents reassigned to another document set

It is no longer possible to add a protected extraction training document to the Classification Set. (1267531)

Extraction training fails after dictionary removed

When a dictionary is removed, all references to that dictionary are also removed. (1267094)

Process batch disables ribbon button

Any Ribbon menu option with a drop-down list of sub-menu options is no longer disabled after the top-level menu option or one of the sub-menu options is selected one time. (1259367)

AZL Image cleanup profiles not working

Image cleanup profiles now behave as expected when used on a zone in an Advanced Zone Locator.

The results of image cleanup as seen in the zone settings is successfully applied before zone recognition is performed. (1255944, 1248195)

Validation object reference error

An object reference error is no longer generated in Validation when user settings are saved. (1247892)

Database lookup field mapping and immediate search

The immediate search can now be configured independently of the "Visible in Search Results" setting.

The "Visible in Search Results" still applies to the database lookup window. (1247785, 1241841)

Two processing windows from extraction training

The Training Extraction behavior has been updated so that the progress window is no longer displayed when the document set is empty. (1246992)

Kofax Transformation Modules 6.2.1 fix pack 13

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 13 (6.2.1.13).

Images cause Recostar to fail

After an update to the Recostar recognition engine, TIFF images are now processed successfully. (1295541)

Validation error on PO_Number confirm

The Validation user module no longer displays errors when there is a line item mismatch. (1283122, 1299668)

Kofax Transformation Modules 6.2.1 fix pack 12

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 12 (6.2.1.12).

Images cause Recostar to fail

After an update to the Recostar recognition engine, TIFF images are now processed successfully. (1291841)

Immediate search breaks in database lookup

There are no longer issues with searches when the Immediate Search option is enabled in the Database Lookup window. (1282314)

Incorrect values displayed at batch close

The correct values are now displayed when a batch is closed in the Validation module. (1278125)

Kofax Transformation Modules 6.2.1 fix pack 11

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 11 (6.2.1.11).

Arabic numbers not correctly recognized

A new scripting property called `DisableFlexiFormsDA` is now available at the `MpsPageRecogProfileFR` Object in script.

This new property improved Arabic recognition and disables the `FlexiFormsDA` setting from FineReader that can interfere with Arabic recognition.

By default, the `DisableFlexiFormsDA` property is set to False.

The following script provides information on how to enable this property through scripting.

```
  Private Sub SetDisableFlexiFormsDA()
     Dim pageProfile As IMpsRecogProfile
     Dim pageProfileFr As MpsPageRecogProfileFR

     Set pageProfile = Project.RecogProfiles.ItemByName("Name of the recognition
profile")
     Set pageProfileFr = pageProfile
     pageProfileFr.DisableFlexiFormsDA = True
  End Sub
```

This code can be called from the Initialize script event as follows:

```
  Private Sub Application_InitializeScript()
   SetDisableFlexiFormsDA()
  End Sub
```

To further improve the recognition for OCR for Arabic Numbers, Along with the above script you also need to set the following settings on the FineReader Page Recognition Profile.

- Select **Correct skew**

- For the **Recognition mode**, select **Fast mode**
- For the **Printer Types**, select **Normal** and ensure that all other printer types are cleared
- From the list of other supported **Languages**, select **Arabic**

(1274708)

Reclassify in script not moving to next document

If a document is reclassified via script and the resulting document is valid, the next document is selected automatically. (1259321)

Merging performance issues in Document Review

There is no longer a performance issue when documents are merged in Document Review. (1257485)

Kofax Transformation Modules 6.2.1 fix pack 10

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 10 (6.2.1.10).

Database lookup immediate search not working

The immediate search can now be configured independently of the Visible in Search Results setting.

The Visible in Search Results still applies to the Database lookup window. (1247784, 1241841)

Folder button missing

Users can now successfully navigate between documents when they keyboard shortcuts. (1219881)

Kofax Transformation Modules 6.2.1 fix pack 9

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 9 (6.2.1.9).

Image cleanup in AZL OMR not working

Image cleanup profiles now behave as expected when used on a zone in an Advanced Zone Locator.

The results of image cleanup as seen in the zone settings is successfully applied before zone recognition is performed. (1248195)

Validation error

An object reference error is no longer generated in Validation when user settings are saved. (1247889)

Kofax Transformation Modules 6.2.1 fix pack 8

This fix pack was superceded by fix pack 9 and fix pack 11.

Kofax Transformation Modules 6.2.1 fix pack 7

The following issues have been resolved in Kofax Transformation Modules 6.2.1 fix pack 7 (6.2.1.7).

Table Locator error

The amount-based table extraction algorithm has been improved so that Fox Exception errors are no longer thrown by table locators. (1229957, 1205615)

Validation hangs because of SecurityBoost issue

If SecurityBoost is enabled, Validation no longer hangs and batch log creation is now successful. (1241379)

Kofax Transformation Modules - Thin Client Server 6.3.0 fix pack 1

The following issues have been resolved in Kofax Transformation Modules - Thin Client Server 6.3.0 fix pack 1 (6.3.0.1).

Manual classification not canceled by script

The Classification combo box is now properly updated when an action is canceled via script. (1272474)

Implement new style for Thin Client Server

The application color has been changed from blue to grey. (1213869)

Kofax Transformation Modules - Thin Client Server 6.2.1 fix pack 12

The following issue has been resolved in Kofax Transformation Modules - Thin Client Server 6.2.1 fix pack 12 (6.2.1.12).

Immediate search breaks in database lookup

There are no longer issues with searches when the Immediate Search option is enabled in the Database Lookup window. (1282314)

Chapter 3

Known issues

The following sections describe known problems, and if available, useful workarounds for Kofax Transformation Modules 6.3.1.

A2iA Document Reader Performance Degradation

During OCR processing, performance degradation occurs when using the A2iA Document Reader recognition engine. (1252327)

Broken forwards compatibility

If you open and save a project in Kofax Transformation Modules 6.3.1, that project can no longer be opened in any previous version of Kofax Transformation Modules.

Clear Thin Client browser history

After an upgrade, certain Thin Client features do not work as expected.

Workaround: Clear the browser history and then restart IIS.

KTM Statistics export connector not installed

If you selected the **Statistics Release** module when you installed Kofax Transformation Modules, the KTM Statistics export connector is not successfully installed. Because of this it is not possible to apply this export connector to a batch class and it is impossible to import and publish a batch class with statistics release.

Workaround: Follow these steps to register the KTM Statistics export connector manually.

1. Open Kofax Capture Administration.
2. From the **Tools** tab, in the **System** group, click **Export Connectors**.
   The **Export Connector Manager** window is displayed.
3. Click **Add**.
   A Windows Explorer window opens.

4. Select **KTMStatisticRelease.inf** and then click **Open**.
   The Add Export Conectors window is displayed.
5. In the Add Export Connectors window, select **KTM Statistics** and then click **Install**.
6. When prompted, click **OK** to confirm that the registration of the KTM Statistics module is complete.
7. In the Add Export Connectors window, click **Close**.
8. In the Export Connector Manger window, click **Close**.
9. Close Kofax Capture.

You can now import and publish batch classes with the statistics export connector.

Named Entity Locator Error

After upgrading to Kofax Transformation Modules 6.3.1 an extraction error may be encountered when testing the Named Entity Locator. (1312052)

Workaround: If you encounter this extraction error, save your project, exit Project Builder, and then restart Project Builder. Also restart Kofax Transformation - Server if it is running. These steps resolve the extraction error so that you should no longer encounter it going forward.

Remove Service Pack 6.3.1 on Windows Server OS

The removal of Kofax Transformation Modules 6.3.1 can fail when using the Windows **Uninstall an update** application. (1313715)

Workaround: Remove Kofax Transformation Modules 6.3.1 manually by following these steps.

Important You must have full Windows administrative privileges to perform the following steps.

You must also have access fo the KofaxTransformationModules-6.3.0.iso file, the original installer program for Kofax Transformation Modules 6.3.0.

1. Mount the ISO file or use a network resource.
2. Terminate all running Kofax Transformation Modules application and services.
3. Unzip the KofaxTransformationModules-6.3.1.zip file.
4. Unpack the KofaxTransformationModules-6.3.1.exe file using the unzip tool. This unpacks the `PatchInstaller.exe` and the KofaxTransformationModules-6.3.1.msp file.
5. Open a Command Prompt as an Administrator and type the following.

   ```
   msiexec /I "<full path>\KTM.msi" MSIPATCHREMOVE="<full path>
   \KofaxTransformationModules-6.3.1.msp" /qb /L*VX "<full path>\KTM-6.3.1-Uninstall.log"
   ```

   Please provide the full path to the mounted MSI file from ISO as the first parameter. Also, provide the full path to the extracted KofaxTransformationModules-6.3.1.msp file as the second parameter.

   Because the MSI runs in silent mode, a log file is recommended as the third parameter.

Note Removing Kofax Transformation Modules 6.3.1 removed the service pack only, and not Kofax Transformation Modules 6.3.0.

Root element error when documents copied via drag & drop

When documents are copied via drag & drop, an error can occur when the project is trained or when the documents are sorted on disk.

Workaround: When presented with this error, press the Ignore or the Ignore All button. Doing so repairs the defective documents. (630605)

Chapter 4

Additional Documentation

The following sections provide additional documentation for new features available in Kofax Transformation Modules 6.3.1, that is not available in existing documentation.

Help updates for Project Builder

The following changes will be made to the *Help for Kofax Transformation Modules - Project Builder* in the next full release of the product.

- **New** indicates a brand new topic that has not appeared in previous versions.
- **Updated** indicates that additions, deletions, and edits exist that do not appear in previous versions.

Updated: Class Details

The following new group and option will be added to this window in the next full release.

Natural Language Processing

This group has the following option:

Language

Select a language from the list so that the Natural language processing engine knows which language to use for extraction. Since this setting is on the class-level, you can configure your classes to use different languages.

If this list is empty you must first install the Natural language processing engine. Information about installing the NLP engine and its languages is found in the *Kofax Transformation Modules* Installation Guide.

If you are using custom entities, the selected language indicates where you should store your custom entity files.

This option is set to the project design language by default.

Extract named entites

When selected, the project executes the extraction of named entities during extraction. These entities can then be referenced in script without the use of a locator.

However, if you want to use the Named Entity Locator and this option is cleared, the extraction of named entities is performed on-demand.

This option is cleared by default.

Stop Salience Engine

Stops the Natural language processing engine. The engine is then restarted on-demand, when needed.

New: Named Entity Locator Properties window

The Named Entity Locator Properties window has the following tabs.

- General tab - Named Entity Locator Properties window
- Regions tab
- Test Results tab

New: General Tab - Named Entity Locator Properties window

This tab enables you to filter the results that are returned by the Natural language processing engine for your needs.

The following options are available on this tab.

Extraction Mode

Select an option to determine how much information is returned. There are two options for this field.

- Simple Field.

  Select this option to return the entity name only.

- Table Field.

  Select this option to return the entity name, confidence, type, and sentiment. When this option is selected, the following additional options are available.

  Table Model

  Select an existing table model from the list or create a new model for this locator by clicking **Table Models**.

  The results returned here come from the entity extraction, so no database table is required.

  Text Column

  From the selected table model, map a column that returns the entity text.

  Confidence Column

  From the selected table model, map a column that returns the entity confidence.

  Entity Type Column

  From the selected table model, map a column that returns the entity type.

  Sentiment Column

  From the selected table, map a column that returns the sentiment of the entity.

Filter Entities By

Use the following settings to refine the extraction results.

This group has the following options:

Entity Type

Select the entity type from the list if you wish to filter the results to a specific entity. If you are using custom entities, type in the name of your custom entity folder.

Confidence Threshold

This is the minimum confidence allowed for an entity to be included in the extraction results.

The value for this option is set to 70 by default.

Sentiment

You can further filter the results by specifying the types of sentiment that are returned. As you select each option, the confidence range is updated for the appropriate selection.

There are several values for this field.

- All. Select this value to return all sentiments found on the document, positive, neutral, and negative.
- Postitive. Select this value to return positive or good sentiments only.
- Neurtral. Select this value to return neutral sentiments only.
- Negative. Select this value to return negative or bad sentiments only.
- Custom. Select this value if you want to customize which sentiments to return. You can control the level using the slider.

  Invert result range

  If selected, the confidence range is reversed so that sentiments that fall outside the selected range are returned only.

New: Sentiment Locator Properties window

The Sentiment Locator properties window has the following tabs.

- Regions tab
- Test Results tab

Updated: Validation sequence

The following description was added after the graphic that explains the sequence of validation and formatting.

1. The first step is to execute any field formatters. If any of these are not successful, the field is marked as invalid and an error message is displayed beside the field.

   The formatters ensure that the extracted text in a field is in the correct format before it is validated.

2. If formatting is successful, all single field validation rules on the fields are executed, one at a time, until one fails.

3. If single field validation is successful, all multi-field rules are retrieved for the document or folder and then each rule is executed, but only if any of its associated fields have changed or require validation. In this case, formatters are also run on any associated fields, if needed.

   If a multi-field validation rule fails then every field associated with the rule is marked as invalid.

New: Manage natural language processing

Kofax Transformation Modules supports natural language processing by extracting named entities and sentiments using the Natural language processing engine.

You can manage natural language processing with one or more of the following.

- Natural language processing engine
- Named Entity Locator
  - Add a table for entity extraction
- Sentiment Locator
- Custom entities
  - Custom entities sample in production
  - Create a custom entity file
  - Edit a custom entity file

New: Natural language processing engine

This engine performs natural language processing for the Named Entity Locator and the Sentiment Locator.

Before you can use this engine and the locators that it supports, a manual installation of the NLP engine is required. For more information on installing the Natural language processing engine, refer to the *Kofax Transformation Modules Installation Guide*.

There is a class-level option called Extract Named Entities. If you are using a script to process the named entities without the use of a locator, this option must be selected. If however, you are using the Named Entity Locator, the extraction of named entities is performed on-demand when the Extract Named Entities class-level option is cleared.

New: Custom entities

Kofax Transformation Modules enables you to define your own custom entities, in any of the supported languages as long as the corresponding language pack is installed. This means that when a specific entity is located on a document, the custom entity file is referenced to determine the extracted result.

For example, your company provides products that are broken down into categories, and these categories each have multiple model numbers. For basic correspondence, the model number is not needed, but the category is always needed. Because of this, you can provide a custom entity file that contains a list of model numbers for each category. When extraction is performed, a model number is recognized, but the category is returned in the extraction results.

Custom entities are stored in a folder structure and entity files must have a specific extension and file format.

If you modify or create a new custom entity file, you must stop the Natural language processing engine to ensure that your changes are available the next time a document is processed. The NLP engine restarts automatically the next time that it is needed.

New: Custom entity folder structure

Custom entities are stored alongside your project files at the following path.

```
\Custom\SalienceData\<language>\salience\entites\<EntityTypeName>\
```

To ensure that you can easily find it, the project path is displayed as a clickable link when you select the Project item from the Project Tree.

You need to create this directory structure if it does not already exist.

There are two variables in this path that must be supplied by you. First, there is the *<language>* folder. This is the two-character code of the language that you are using. For example, the following are the languages that are supported by the Natural language processing engine along with the two-character codes that you must use to name the folder.

- German - `de`
- English - `en`
- Spanish - `es`
- French - `fr`
- Italian - `it`
- Japanese - `ja`
- Korean - `ko`
- Dutch - `nl`
- Portuguese - `pt`
- Romanian - `ro`
- Chinese - `zh`

The second variable is the name of the folder under the `entities` folder. The name of this folder is what determines the Entity Type that is returned by the NLP engine. The entity type is visible when using a table field and the Named Entity Locator, or when viewing a test document in the XDoc Browser.

For example, if you have a custom entity in a folder called Cars, the NLP engine is able to recognize that this is a noun and that it is in plural format. Because of this, the returned Entity Type in this example is actually "Car." If the folder name is already singular, or it is not recognized as plural, the folder name is returned instead, as long as the entity `.cdl` file is in the correct format.

New: Custom entity file format

The entity files use the `.cdl` file extension and must be formatted using the following tab delimited structure.

```
Entity Search Text<tab>Entity Label<tab>Normalized Entity Name
```

Entity Label and Normalized Entity Name are optional, but recommended.

If the file contains values for the Normalized Entity Name, this is what is returned by the Natural language processing engine.

For example, your `.cdl` file contains the following content.

```
Dog <empty> Canine
Cat <empty> Feline
```

When extraction is performed, the NLP engine returns either Canine or Feline rather than Dog or Cat.

Note The Entity Label value provided in the .cdl file is not currently supported. This is why that value is blank in the above example.

New: Custom entities sample in production

You have a custom entity file that has the following content using the recognized format for an entity file.

```
Ford F150 Car Ford F-series truck
Ford F250 Car Ford F-series truck
Ford F350 Car Ford F-series truck
Honda CBR 125 Motorcycle Honda CBR motorbike
Honda CBR 650 Motorcycle Honda CBR motorbike
```

This custom entity file is used to locate all occurrences of the specified entities and returns the normalized entity name as the extracted value.

Note The Entity Label is currently not used by Kofax Transformation Modules.

The following content from a document is processed.

```
I'm driving a Honda CBR 125, and would like to buy a Ford F250.
```

Using this custom entities file, two results are returned when this document is processed.

1. Ford F-series truck
2. Honda CBR motorbike

New: Create a custom entity

If you have a Named Entity Locator that requires a new entity file, this file is not visible to the Natural language processing engine until after the engine is restarted. Because of this, it is necessary to stop the NLP engine after the entity file is created to ensure that your changes are available the next time the NLP engine is run. The NLP engine restarts on-demand so that it does not use server resources when not in use.

You can create a new custom entity file and then stop the Natural language processing engine by following these steps:

1. On the **Project Tree**, click on the **Project** node.
   The properties of the project are displayed, including a link to the location of the project file.
2. Click on the **Project path**.
   A Windows Explorer window opens to the location of your project.
3. If it does not already exist, create the following folder structure.
   `\Custom\SalienceData\<language>\salience\entites\`
   Where *<language>* is the two-character code for the language that you are using.

4. Inside the `entities` folder, create a new folder with the name of the type of entities you are creating.

   The name of the folder, or a singular version of this name, if appropriate, is returned as the Entity Type.
5. Inside the newly created entity type folder, create a file with the `.cdl` extension.

   The name of this file is not used by the NLP engine, but ensure the name is useful and descriptive.
6. Add content to the new `.cdl` file in the following format.

   ```
   Entity Search Text<tab>Entity Type<tab>Normalized Entity Name
   ```
7. Save and close the file.
8. Navigate to and select the class in the **Project Tree** where your **Named Entity Locator** is located.

   The details for that class are displayed.
9. Scroll down and find the **Natural Language Processing** group and then click **Stop Salience Engine**.

   The Natural language processing engine stops. The next time the NLP engine is needed, it starts on-demand and your new custom entity file is now visible.

New: Edit a custom entity

If you need to make changes to an existing custom entity file, those change are not visible until after you restart the Natural language processing engine.

You can edit a custom entity file and then stop the Salience engine by following these steps:

1. On the **Project Tree**, click on the **Project** node.

   The properties of the project are displayed, including a link to the location of the project file.
2. Click on the **Project path**.

   A Windows Explorer window opens to the location of your project.
3. Navigate to the following path where your entity files are stored.

   `\Custom\SalienceData\`***<language>***`\salience\entites\<EntityTypeName>`
4. Open the `.cdl` entity file that needs modified.
5. Make the necessary changes and then save and close the `.cdl` file.
6. Navigate to and select the class in the **Project Tree** where your Named Entity Locator is located.

   The details for that class are displayed.
7. Scroll down and find the **Natural Language Processing** group and then click **Stop Salience Engine**.

   The Natural language processing engine stops. The next time the NLP engine is needed, it starts on-demand and your modifications to your custom entity file are now visible.

New: Custom sentiments

By default, the Natural language processing engine provides a dictionary that contains keywords or phrases and their sentiment value between -1 and 1, indicating whether they are positive, neutral, or negative. In order for a word or phrase to be included in the sentiment calculation for a document, it must be included in the dictionary file. There is a dictionary file for each of the installed languages.

You must store custom sentiment dictionaries alongside your project in the following location.

```
\Custom\SalienceData\<language>\salience\sentiment\general.hsd
```

Where `<language>` is the two-character code of the language that you are using. For example, the following are the languages that are supported by the Natural language processing engine along with the two-character codes that you must use to name the folder.

- German - `de`
- English - `en`
- Spanish - `es`
- French - `fr`
- Italian - `it`
- Japanese - `ja`
- Korean - `ko`
- Dutch - `nl`
- Portuguese - `pt`
- Romanian - `ro`
- Chinese - `zh`

If a commonly used word or phrase is not included in the dictionary, and the sentiment value that is generated does not match your expectations, you can create a local copy of the dictionary `general.hsd` file and edit it as needed.

Important All entries in the copied `general.hsd` are in alphabetic order. The format is `Word or Phrase<tab>Value between -1 and 1.`.

Additionally, if a word or phrase is included in the dictionary, but its value does not match how that word or phrase used in your local language, you can modify the value as needed. For example, the phrase "kiling it" would produce a very negative sentiment because the verb is recognized instead of the intended adjective. Similar results are found with the phrase "da bomb". If you encounter these phrases, or similar phrases frequently in your documents, consider adding them to your `general.hsd` file to influence the default negative sentiment.

New: Add custom sentiments

If you want to add a word or phrase to the sentiment dictionary, or modify the sentiment value of an existing word or phrase, you can do this by following these steps:

1. On the **Project Tree**, click on the **Project** node.
   The properties of the project are displayed, including a link to the location of the project file.
2. Click on the **Project path**.
   A Windows Explorer window opens to the location of your project.
3. If it does not already exist, create the following folder structure.
   `\Custom\SalienceData\`***<language>***`\salience\sentiment\`
   Where *<language>* is the two-character code for the language that you are using.
4. Copy `\Program Files (x86)\Common Files\Kofax\Salience6.4\`***<language>***`\salience\sentiment\general.hsd` to the path in the previous step.
5. Edit the `general.hsd` file as needed by adding the new word or phrase on a new line, in the correct alphabetic order, using the following format.
   `Word or Phrase<tab>Value between -1 and 1.`
6. Save the dictionary file.

7. Navigate to the **Class Details** and click **Stop Salience Engine**.
   This stops the NLP engine so that the next time it is needed, it starts on-demand and loads your updated `general.hsd` file.
8. Test some documents to ensure that the resulting sentiment values match your changes.

Updated: Database Locator

The following sentence was added to the Database Locator topic and will appear in the next full release.

If the data you want to use is in a relational database, you must import that data into a fuzzy database before you can use it in a database locator.

New: Named Entity Locator

The Named Entity Locator is used to assign extracted entities to fields.

Named entities are text grouping such as a people, places, organizations, roles, times, amounts, etc. These named entities are found in unstructured, natural language text like sentences found in emails, documents, or even a report, to name a few.

Because this type of data is difficult to extract with other locators unless you have a known set of criteria to search for, such as a database with people names, this is where the Named Entity Locator comes in. It is able to extract data based on its grammatical placement in a document.

If you need to extract more than one type of entity, a separate Named Entity Locator is required for each entity type.

Once the entities are extracted, a customized script is needed to interpret the entities. For more information on scripting, refer to the *Help for Kofax Transformation Modules Scripting*.

When configuring a Named Entity Locator, the biggest decision is how much information you want. If you are interested in the entity name only, use a simple field, as this is the only information returned. If you want to know the entity name, entity type, confidence, and sentiment of a document, use a table field.

The Named Entity Locator Properties window has the following tabs that enable you to configure the extracted entities.

- General tab - Named Entity Locator Properties window
- Regions tab
- Test Results tab

New: Sentiment Locator

The Sentiment Locator uses the built-in functionality of the Natural language processing engine to extract text sentiments from a document. This means that the Sentiment Locator is able to determine the overall mood, or sentiment of a document based on the words and phrases found on a document.

For example, if your company has several products, and each have online ratings, you can use the Sentiment Locator to determine whether your users are providing good and positive reviews, or bad and negative reviews.

This locator is mapped to a simple field and the extracted result of the locator is any value between -1.00 and 1.00 with the following ranges.

- Positive - `0.12` to `1.00`
- Neutral - `0.025` to `0.11`
- Negative - `0.026` to `-1.00`

For example, a result of `-0.75` indicates that the processed document contains a majority of bad or negative comments. A result of `0.8` indicates that the processed document contains a majority of good or positive comments, and a result of `0.1` indicates that there is about an equal ratio of both positive and negative comments.

Because this locator searches for sentiments automatically there are no configurable settings for this locator. You can however, narrow the area on a document or page where the locator is executed. You can do this by adding regions to a document or page so specific areas are included or excluded from the locator.

For example, if your document contains a header that includes an address and a salutation, these areas do not typically contain data related to the overall mood or sentiment of a document. As a result, you can use a region to exclude the top 1/3 of the first page of a document.

The Sentiment Locator Properties window has the following tabs.

- Regions tab
- Test Results tab

Help updates for XDoc Browser

The following changes will be made to the *Help for Kofax Transformation - - XDoc Browser* in the next full release of the product.

- **New** indicates a brand new topic that has not appeared in previous versions.
- **Updated** indicates that additions, deletions, and edits exist that do not appear in previous versions.

New: Nlp

When this item is selected in the XDoc Browser Object Tree, it displays all Entities and their Mentions. Each Entity has a list of Mentions that show where that entity is found in a document.

Updates for the Installation Guide

The following changes will be made to the *Kofax Transformation* - Installation Guide in the next full release of the product.

- **New** indicates a brand new topic that has not appeared in previous versions.
- **Updated** indicates that additions, deletions, and edits exist that do not appear in previous versions.

New: Install Natural language processing engine

Before you can use either the Named Entity Locator or the Sentiment Locator it is necessary to install the Natural language processing engine. This engine performs natural language processing and is able to extract items such as named entities and sentiments.

For more information on the Natural language processing engine and its supported locators, refer to the *Help for Kofax Transformation Modules - Project Builder*.

Download the Natural language processing engine from the Kofax Edelivery site.

You can install the Natural language processing engine by following these steps:

1. Extract the downloaded zip file.
2. Navigate to the `MicroPackage` folder in the extracted files and double-click on one of the following `.MSP` files, depending on what languages you are supporting.
   - **KofaxTransformation_SalienceV6.4.0_LanguageBundle_western-default**.
     Run to install English, Spanish, Portuguese, French, and German support.
   - **KofaxTransformation_SalienceV6.4.0_LanguageBundle_western-extended**.
     Run to install Italian, Romanian, and Dutch support.
   - **KofaxTransformation_SalienceV6.4.0_LanguageBundle_extended**.
     Run this to install Japanese, Chinese, and Korean support.

   A Windows Installer window appears and then installs the Kofax NLP (Natural language processing engine) automatically.

   The installer window closes when the installation is complete.
3. Optionally, double-click on another installer if you want to support additional languages.