Notice that online learning is Enabled. The server will use the dynamic knowledge base during extraction. And automatic training after validation has been enabled.

But…there is a limit to how many training documents are stored in the dynamic specific knowledge bases. 2000 is an default number. Making it larger may slow performance. So…when the limit is reached, no new training documents; no new learning. Maintenance is required.
## Specific Knowledge Bases and Confidence Scores

<table>
<thead>
<tr>
<th>Number of Specific Samples</th>
<th>Resulting Confidence Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50%</td>
</tr>
<tr>
<td>2</td>
<td>85%</td>
</tr>
<tr>
<td>3</td>
<td>95%</td>
</tr>
<tr>
<td>4</td>
<td>100%</td>
</tr>
</tbody>
</table>

The numbers represent the number of samples per given layout. With one document, the results are recognized correctly, but with a low confidence score, colored red in Validation. Trainable fields require confirmation (the [Enter] key). So a second sample is needed. Now the confidence scores are within the default 80% minimum required by the field properties and are colored green, which means they no longer require Validation.
Knowledge Bases – Generic and Specific

Transformation Designer

Release

Import

Production

Training Documents

Dynamic Specific KBs

Published Generic Bs

Published Specific KBs

Server

Validation

Knowledge Base Learning Server

KOFAX
Specific Knowledge Base - Maintenance

• Time interval for the maintenance

Assuming that there are no Specific Knowledge Bases available at day one of production, the document will be extracted by GKBs and rules-based methods. Documents with trainable fields that are read at a low confidence level or have their data changed in Validation are marked for specific learning. This means that the Specific New Samples count will increase rapidly. The system trains at least three documents per layout for Specific Knowledge Base extraction.

A user with 50,000 vendors in their vendor database could theoretically have 50,000 x 3 Specific Samples per vendor. Therefore, it is fair to assume that in the first few weeks of production the number of Specific New Samples will reach the maximum very quickly. Once the frequent vendors, the vendors that the user does the most business with, have been learned by the system, the number of Specific New Samples will increase at much lower rate, which means that the interval between the Knowledge Base maintenance is less frequent.
1. First download (required when you are approaching the first 2000 samples) to a physical folder called Week1. This cleans out the "New Samples" folder in the TotalAgility DB.

2. Select all documents and click Import. This will create a new Extraction set called "Week1".

3. Create an Extraction set called SpecificKnowledgeBase.

4. Move all Week1 documents to the SpecificKnowledgeBase Extraction set.

5. Train for extraction, SpecificKnowledgeBase to identify conflicts.

6. Resolve conflicts. Train until all conflicts are resolved.

7. For each trainable locator, build a Specific Knowledge Base from SpecificKnowledgeBase. Select the knowledge base for use.

8. For SpecificKnowledgeBase document set, Exclude Subset from Training. (No longer needed because knowledge bases now include all of the training information).

Knowledge Base Maintenance (cont.)

1. Subsequent Downloads: Continue job processing. When you near 2000 new samples…

2. Import Specific "New Samples" into the Extraction group.

3. [Select All] new samples in the newly created set (Week 2, etc.). Move them into “SpecificKnowledgeBase" to add them to the documents already there. Re-enable it for training by clicking on "Include Subset for Training."

4. Train for extraction, " SpecificKnowledgeBase to identify conflicts.

5. Resolve conflicts. Retrain & resolve until all conflicts are resolved.

6. For each trainable locator, build a new Specific Knowledge Base from " SpecificKnowledgeBase." Select the knowledge base for use. De-select the knowledge base you created last time. All of its training in included in the newly-created knowledge bases.

7. For " SpecificKnowledgeBase " document set, "Exclude Subset from Training." (No longer needed because knowledge bases now include all of the training information). Remove empty document sets. Repeat these steps when necessary.

8. Train and release the Extraction Group – Re-release any associated Classification group
A conflict is created when an operator selects a result for the same trainable field on more than one document with the same layout at different coordinates on the document.

Training for extraction will automatically identify such conflicts.

Select class and extraction training set. Then click on the Extraction button to train.
Click on the [Resolve Conflicts] button to begin conflict resolution. As you proceed, resolve the conflict and then use the navigation button on the conflict resolution windows to proceed to the next document.
Resolve all conflicts. Then retrain to check for any new conflicts created. Once retraining reveals no new conflicts, you're ready to build your knowledge bases.
Display the properties of each trainable locator and create a new knowledge base. Remove any old Specific knowledge bases.

Select the "SpecificKnowledgeBase" training folder. Then click [Create].
SKB Created - Select Knowledge Base for Use

Check the checkbox to select for use.
Exclude All Specific Samples – No Longer Needed

Remember to retrain and release your Extraction Group.
Sometimes organizations want to password-protect knowledge bases (eg. Kofax partners selling their knowledge bases to end-user customers). Password-protected knowledge bases are specific to the serial number of one Kofax License number and require activation to use…

When creating the knowledge base, you must provide a password.
Generating Activation Code

To use the created knowledge base, an activation code (specific to the license) will need to be created...
Enter the original password used to create the knowledge base and the serial number of the Kofax license for the system on which you want it activated.

Both the serial number and the activation code will be required to activate the knowledge base.
To activate the knowledge base on the new machine:
1. Copy the knowledge base file to the new machine.
2. Import the knowledge base.
3. Click on [Insert].
4. Enter the serial number of the Kofax license.
5. Enter the activation code provided.
6. Click [OK]
The license is activated. Check the checkbox to use. Repeat this process for all other knowledge bases.