ABBYY Headquarters

P.O. Box 20, Moscow
127273, Russia
Tel.: +7 (495) 783 3700
Fax: +7 (495) 783 2663
office@abbyy.com
www.ABBYY.com

ABBYY FlexiCapture 12 Release 2
Release Notes

Table of Contents

Introduction

About this document

This document describes the features that have been implemented in ABBYY FlexiCapture 12 Release 2.

About the product

ABBYY FlexiCapture 12g is the latest version of ABBYY FlexiCapture, a highly scalable and flexible data capture platform for creating regional and vertical data capture solutions.

Installing ABBYY FlexiCapture 12

ABBYY FlexiCapture 12 can be installed on the same computer where ABBYY FlexiCapture 10 or 11 is already installed. Please refer to the ABBYY FlexiCapture 12 System Administrator's Guide for installation details.

Upgrading from previous versions

ABBYY FlexiCapture 12 can be installed on the same computer where ABBYY FlexiCapture 10 or 11 is already installed. You cannot use the setup wizard to upgrade your ABBYY FlexiCapture 10 or 11 to version 12.

ABBYY FlexiCapture 12 Release 2 can be used as an upgrade if you have ABBYY FlexiCapture 12 Release 1 (with any updates) installed on your computer. Release 1 will be removed and Release 2 will be installed instead.

You can use ABBYY FlexiCapture and ABBYY FlexiLayout projects and FlexiLayouts created in earlier versions of the program. If you have documents already loaded into the system, we recommend that you first complete the processing of these documents and only then migrate to ABBYY FlexiCapture 12. When you open an ABBYY FlexiCapture or ABBYY FlexiLayout project created in an earlier version of the program, it will be converted to the ABBYY FlexiCapture 12 format. Once a project is converted to the new format, it can no longer be opened in an earlier version.

A detailed description of the upgrade procedure can be found in the System Administrator's Guide.

OCR Technologies

ABBYY FlexiCapture 12 Release 2 uses a new version of OCR Technologies (v. 16), delivering a better overall quality of OCR compared to FlexiCapture 12 Release 1 and earlier versions.

- Overall OCR quality has been improved. Slight slowdowns may be detected, mostly on Arabic texts, due to the increased number of detected lines.
- Barcode detection quality has been improved by 3%, with a 4% loss in speed. Detection of MaxiCode barcodes has been significantly improved.

Important! Please note that due to changes in the recognition technology, the FlexiLayout matching process in the new version can be slightly different. If a project that you created in an earlier version was designed to accommodate for certain OCR defects, you may now need to change its logic. If you upgrade a project created in an earlier version, please try matching your FlexiLayout on sample documents first.

Licensing

For ABBYY FlexiCapture 12 Release 2, you need a serial number generated specifically for this version 12. Serial numbers generated for ABBYY FlexiCapture 12 Release 1 (with any update) can also be used for Release 2.

Serial numbers for previous versions of ABBYY FlexiCapture cannot be used for ABBYY FlexiCapture 12.

Localization languages

ABBYY FlexiCapture 12 interface is translated into the following languages:

English, Russian, German, French, Spanish, Korean, Japanese, Czech, Serbian (Latin), Portuguese (Brazil), Polish, Chinese Simplified, Hungarian.

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 6	**1299/29**	**12.0.2.1420**	**29512**	**16.1.251.84**	**2018.12.28**
Release 2 Update 5	1299/28	12.0.2.1398	27056	16.1.251.75	2018.12.05
Release 2 Update 4	1299/27	12.0.2.1387	25159	16.1.251.75	2018.11.22
Release 2 Update 3	1299/26	12.0.2.1370	22855	16.1.251.68	2018.10.30
Release 2 Update 2	1299/25	12.0.2.1363	21603	16.1.251.66	2018.10.19
Release 2 Update 1	1299/24	12.0.2.1356	20560	16.1.251.66	2018.10.05
Release 2	1299/22	12.0.2.1194	16090	16.1.251.47	2018.08.03

Key Features

1. New image enhancement functionality
2. Improvements in advanced classification
3. Unstructured text processing
4. Oracle and Microsoft Azure SQL Database support
5. Web Capture Station added
6. New HTML5 functionality
7. HTML5 Web Stations Customization
8. Reverse matching
9. Platform and station improvements
10. Security issues addressed

New Features

1. Image enhancement

The new image enhancement and quality assessment features can be used to improve poor-quality images. A self-service mode for sending documents to FlexiCapture is provided. In this mode, document images can be received form various sources (e.g. photos from mobile devices) and enhanced to prepare them for OCR.

Images can be enhanced manually, by selecting necessary tools or profiles in the UI, or automatically, using predefined or previously created profiles.

1.1. Image enhancement operations

There is a new pane on the Scanning Station that allows applying necessary enhancement operations to each image. The pane is opened by clicking the  button on the toolbar or by selecting **Page > Enhance Image** command from the menu.

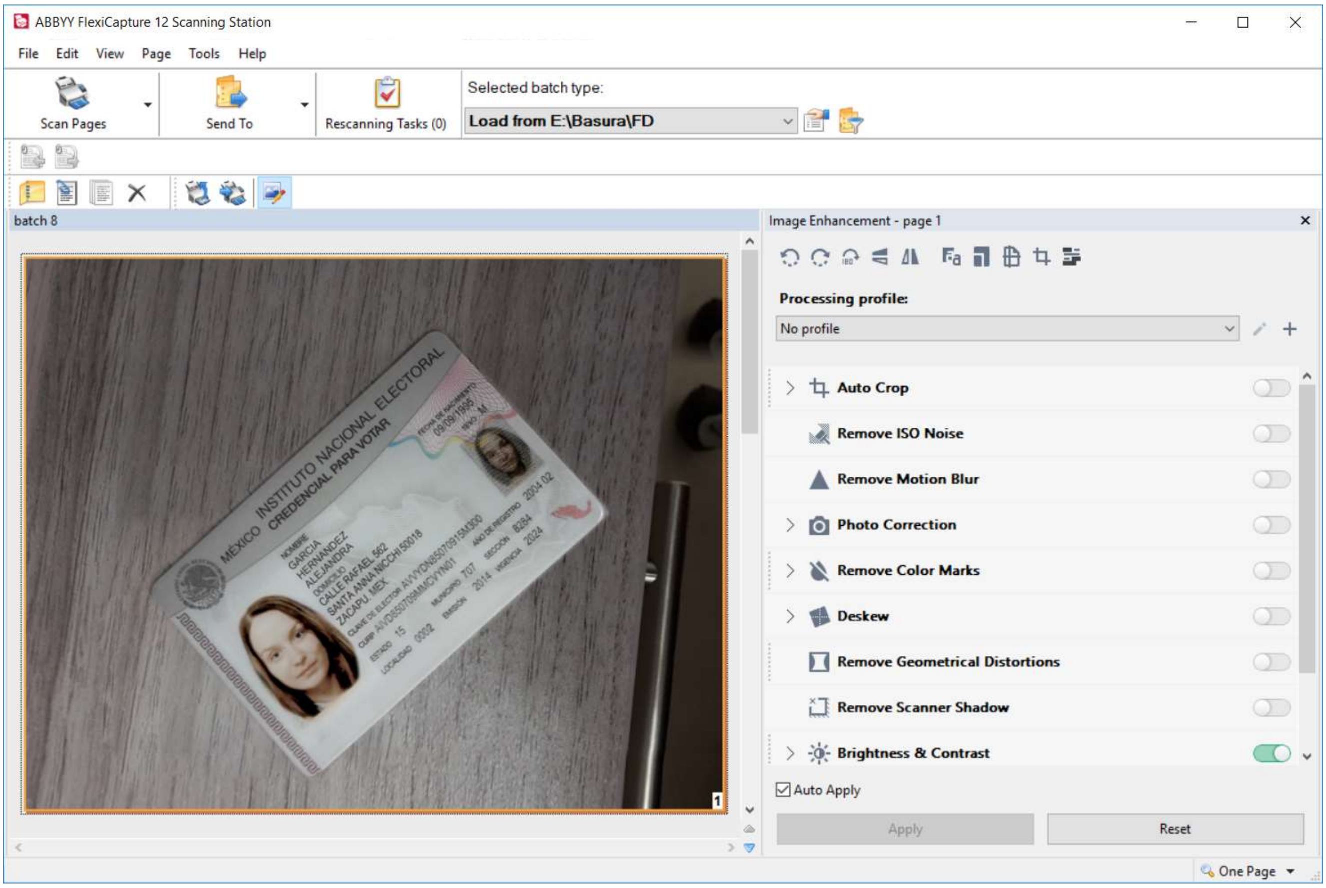
The pane has two sections:

1. The operations displayed at the top of the pane can be applied separately.
2. The operations displayed below the processing profile name can be used to create a sequence. The operations in a sequence will be applied to an image consecutively top down. To include an operation into a sequence, specify its parameters and turn in it on. To change the order of operations in a sequence, you can move them about using drag-and-drop or by pressing a Shift + keyboard arrows. Please note that only the operations that have a dot on the left can be moved.

Note: Normally, you will not need to change the default order of operations, but in some cases changes may be required. For example, on some images, color background lines may have to be removed after changing the contrast and brightness, in which case you need to put **Brightness & Contrast** above **Remove Color Marks**.

You can also use profiles for image enhancement. See *1.2 Image enhancement profiles* for more information.

Profiles and any of the operations displayed at the top of the toolbar can also be applied to multiple images, whereas a sequence of operations can only be applied to a single image.



1.2. Image enhancement profiles

Profiles are named sequences of image enhancement operations that you save by clicking the buttons on the right of the profile name. You can create your own profiles or use one of the two predefined profiles, **Scan** and **Photo**. Profiles can be applied manually or automatically.

To manually apply a predefined or a previously created profile to an image, simply select it from the drop-down list. The operations included in the profile will be applied to the selected image.

You can also apply profiles automatically. See *1.4 Applying profiles automatically* for more information.

Any previously created profile (with the exception of the predefined **Scan** and **Photo** profiles) can be modified and saved with the same or a different name. The predefined **Scan** and **Photo** profiles cannot be modified directly, but you can create a copy of a predefined profile and modify the copy.

Any of the following four operations can be added into a profile: **Split**, **Change resolution**, **Rotate**, and **Flip**. These operations become available for inclusion when you start editing a profile.

To manage your image processing profiles, click **Tools > Processing Profiles**. This opens a list of your current profiles where you delete profiles, export profiles to XML, or import profiles created on other stations.

Profiles can be created and managed on the Project Setup Station or on the Administrator Station. On the Administrator Station, click **Project** > **Project Properties** and then on the **Image Enhancement Profiles** tab click the **Edit** button. The available operations and the UI are identical to those on the Scanning Station.

1.3. Importing and exporting profiles

Image processing profiles can be connected to image sources on other Scanning Stations, on the Project Setup Station, or on the Administrator Station. To use a profile on another station, you need to export it to XML and then import the XML file on the station where you want to use the profile.

To export or import a profile:

- On the Scanning Station, click **Tools > Processing Profiles** and use the **Export…** or **Import…** buttons.
- On the Project Setup Station or the Administrator Station, click **Project** > **Project Properties** and use the **Export…** or **Import…** buttons on the **Image Enhancement Profiles** tab.

1.4. Applying profiles automatically

Profiles can be applied manually as described in 1.2 above, but they are primarily intended for automatic image processing.

Profiles can be applied automatically after scanning images on a Scanning Station, after adding images from the Hot Folder or from the FlexiCapture Mobile client, or after obtaining images from other sources. To apply a profile automatically, select it in the image processing settings:

- For a batch type on the Scanning Station, click **Tools > Batch Types**, select the desired batch type, click **Edit**, and on the **Image Processing** tab select **Apply processing profile**.
- For a batch type on the Project Setup Station or on the Administrator Station, click **Project > Batch Types**, select the desired batch type, click **Edit**, and on the **Image Processing** tab select **Use image enhancement profile**.
- For a import profile on the Project Setup Station or on the Administrator Station, click **Project > Image Import Profiles**, select the desired import profile, click **Edit**, and on the **Image Processing** tab select **Use image enhancement profile**.
- When loading images manually, click the **Load images** command, then click the **Image Processing Settings**… button and select **Use image enhancement profile.**

On the Scanning Station, you can use only one image enhancement profile, because this station is only intended for scanning images. Elsewhere, you may have scans and photos all in one flow, in which case you may want to specify different profiles for scans and photos. The program will check the metadata of an imported file and apply the appropriate profile. If you want to apply one profile both to scans and photos, select the desired profile from the first drop-down list and make sure that the **Use special profile for photos** option is not selected.

1.5. New image processing operations

Several new image processing operations have been added onto the image enhancement pane of the Scanning Station. Several image processing operations have also been added on the Project Setup Station and on the Administrator Station, so that they can be used without creating a profile. These new operations are: **Remove all color marks**, **Correct resolution of photos**, **Autocrop**, **Reduce ISO noise**, and **Whiten background**.

To select any of the new preprocessing operations, select the **Use base image enhancement** option on the **Image Processing** tab and click the **Settings** button.

1.6. Going back to the original image

You may need to go back to the original image if the preprocessing settings result in loss of information. For example, in rare cases the autocrop operation may crop only one page of a two-page passport image.

On the Scanning Station, you can always go back to the original image by clicking **Back to original image** on the shortcut menu of the page.

At the verification stage, you will need to save the original image before you can go back to it. By default, original images are not saved because this would significantly increase the size of the batch. To save original images, select the **Save original image** option on the **Image Processing** tab.

To go back the original image at the verification stage, click **Edit image > Revert to original** on the shortcut menu of the image.

1.7. New scripting methods

New scripting methods have been added to apply profiles from scripts, to apply the new operations, and to check if image quality is good enough for OCR.

The following new methods have been added to the IPictureObject interface:

bool IsSuitableForOcr(); - detects if the image is good enough for OCR

```
float CalcSuitabilityForOCR(); - returns a confidence level for how good the image is for OCR
```

The new property IsPhoto detects if the image is a photo (using EXIF metadata).

Example of how to use the property in a script. The script runs on pressing a custom button and writes the quality evaluation results into the page comment.

```
foreach( IDocument doc in MainWindow.Selection.Documents ) {
  MainWindow.TaskWindow.OpenDocument( doc );
  foreach( IPage page in doc.Pages ) {
    IPictureObject picture = page.Picture;
    double score = (double)picture.CalcSuitabilityForOCR();
    bool verdict = (bool)picture.IsSuitableForOcr();
    page.Comment = score.ToString( "0.0000" ) + " / " + verdict.ToString();;
  }
}
```

The next example shows how to apply a profile to a page depending on the class assigned to this page after classification.

```
IBatchTypeClassifierResult classificationResult = BatchTypeClassifier.ClassifyPage(Page, null);
FCTools.ShowMessage("Class name: " + classificationResult.Classes[0].Name);
string str;
str = classificationResult.Classes[0].Name;

if(str.Contains("HALLOWEEN Form"))
{
  //apply Photo 1 profile if class is "HALLOWEEN Form"
  Page.ApplyImageEnhancementProfile("Photo 1");
}
else
{
  //apply profile Photo 2 otherwise
  Page.ApplyImageEnhancementProfile("Photo 2");
}
```

2. Improvements in advanced classification

2.1. Classification confidence and verification of classification results

Pages and documents that have been classified with a low level of confidence can be submitted for manual verification and highlighted in the GUI. This facilitates scenarios where classification is used for identifying document types, sorting documentsor assembling documents.

As a result of classification, each page is assigned one of the possible classes or remains unclassified. For each classified page, a classification confidence level is provided. You can use scripts to determine if the operator should verify a page or a document classification result. The determination should be made for each page, based on its classification result and classification confidence level. If a document contains at least one page that requires verification, the whole document will have to be verified.

Generally, when configuring the verification of classification results, you will want to achieve the following:

- Eliminate false positives by verifying all of them
- Increase recall when you do not need to make decisions based on assigned classes and just want to confirm suggested classes automatically.

It is important to understand that the only way to minimize false positives is to train your model for high precision and apply some business rules after the classification step to capture false positives. Even a well-trained model will still produce false positives, because there is no technology that will provide 100% quality. If there are decisions that can be made at the classification step (e.g. based on page features), they can be incorporated in your high-precision model, or you can add them as a rule-based layer to the model.

How to set up a script that will determine if page or document requires verification

When you classify a page using a classification script, you not only get the final decision about the class of that page or document, but also have access to the classification probabilities that the program calculated when making the decision. For example:

Class A, 90%

Class B, 30%

…

Class N, 1%

You can use this information to determine the class of a page and if a page requires verification. To specify that a page requires verification, set Page.IsResultClassConfident = false (the default value is ‘**true’**). If a document has at least one page that requires verification, then the entire document must be verified.

The typical way to determine if a page requires verification is to consider the best class and it’s probability (this probability can be termed “absolute threshold”) and, perhaps, the second most probable class and the distance between the probabilities of the first and the second most probable classes (this metric can be termed “relative threshold”). Suppose you have a page that is classified as ‘Invoice’ with a probability of 90% and as a ‘Contract’ with a probability of 10%. In this case, the absolute threshold would be 90% and the relative threshold would be 80%. Then it is most probable that the page is indeed an invoice and no verification is required. Now suppose that besides the ‘Invoice’ and ‘Contract’ classes your page is also classified as ‘Purchase Order’ with a probability of 50%. In this case the absolute threshold would be 90% and the relative threshold would be 40% and it is best to verify the class. This approach based on absolute and relative thresholds is easy to understand and script.

Example:

```
IBatchTypeClassifierResult classificationResult = BatchTypeClassifier.ClassifyPage(Page, null);

int absThreshold = 0;

int relThreshold = 100;

if (classificationResult.Classes.Count > 0) absThreshold = classificationResult.Classes[0].Confidence;

if (classificationResult.Classes.Count > 1) relThreshold = absThreshold - classificationResult.Classes[1].Confidence;

if (absThreshold > 80) Page.ResultClassName = classificationResult.Classes[0].Name;

if (relThreshold<50) Page.IsResultClassConfident = false;
```

The real difficulty is that in order to select good thresholds you have to experiment a lot. At some point you may find it useful to use different threshold values depending on the first most probable class and on the second most probable class.

There is yet another approach to class verification. A classification script allows you to combine different classification models to classify one page (see “Combining classification models for hierarchical classification”). The idea is to train two classification models with different recall and precision priorities for one classification task.

Let’s assume that you have trained a classification model prioritizing precision, and the precision and recall it now delivers are respectively Pp = 99% and Rp = 70%. This means that you are able to classify properly 70% of pages, leaving the rest 30% either unknown or with the wrong classes assigned. The 99% precision value means that among all the pages that have a class assigned to them, 1% will have the wrong class assigned.

The first idea that naturally presents itself is to find a way to verify this 1% of false positives. The bad news is that the only way to do this is to apply some business logic after the classification step. For example, you could extract document fields and validate them using rules, or you could check the integrity of the batch to see if some pages (and, consequently, documents) have been classified incorrectly. If you had some page-dependent rules, you could *build them into the classification mode itself. But you couldn't possibly improve this result by only experimenting with thresholds, because (a) it is impossible to achieve 100% precision and (b) all the available tricks to maximize precision (i.e. minimize false positives) have already been utilized when training the model, including the selection of appropriate thresholds. Conclusion: you either have trust this model and accept this 100% - Pp = 1% of false positives (and maybe think of how we can address the issue in the processing steps that follow) or you have to verify 100% of the classified pages.*

What you really can improve in this case is recall. Let's assume that you have trained another classification model prioritizing recall, and the recall and precision it now delivers are respectively Rr = 97% and Pr = 90%. The first model will leave 100% - Rp = 30% of the pages unclassified. If you classify the same pages using the high-recall model, only 3% of the pages will be left unclassified. Keeping in mind that the high-recall model is less precise, we can suggest the following scheme:

- *If both models agree on a class, assign this class to the page and consider it as confidently assigned.*
- *If the high-precision model can't classify a page whereas the high-recall model produces some class, select this class and require verification.*
- *If both models can't classify a page , leave the page unclassified*
- *If the two models assign different classes to a page, trust the high-precision model but require verification.*

It is very likely that we will finally get the following result from the combination of two models: 70% (the recall of the high-precision model) of the pages will be classified and require no verification, while 1% will be false positives (100% minus the precision metric of the high-precision model). If we apply only the high-precision model, we will get 30% of unknown pages and most likely we will have to classify them manually. And if we add the high-recall model, nine out of ten pages (since 90% is the precision of the high-recall model) will be classified properly within 29% of these documents and only 1% will be left unknown. This is a huge help to alleviate the manual classification of documents that have not been classified by the high-precision model.

A document where at least one page requires verification will be sent to the "Verification" and "Document Assembly Check" stages, if these stages are enabled. By using a special stage entry condition ("Has pages that need class validation") in the advanced workflow, you can create a separate stage just for verification of classification or adjust the workflow to your use case.

Low-confidence pages and documents containing such pages will be highlighted in red. The operator can change the document type (by clicking the "Match Document Definition" button) or confirm the current type (by clicking the "Confirm Document Definition" button). Either action will remove the red highlighting from the document.

Sometimes the fact that the section successfully matches the pages means that the page class has been identified correctly and no verification is required. Usually, this depends on the section. For example, if your FlexiLayout has a strong identifier and footer, you can choose the "Automatically confirm section type when matched" option in the section properties.

2.2 Continuous improvement based on supervised learning

You can enable auto-learning to improve document classification and adapt the classifier to changes in your business processes. With auto-learning configured in your Document Definition, the program will learn from corrections made the verification operator. The operator may need to correct the result of automatic classification when a document is put in the wrong class or when the programs fail to assign a class altogether. Then the job of the verification operator is to assign the correct class to the misclassified document. Any documents so corrected will be placed in a batch named "Classification Training Batch." When the number of corrected documents received from the verification operator reaches 30% of the documents in the Classification Training Batch, auto-learning will start automatically.

As a result, a new classification model will be created. This new model will be compared against the old model and the winning model will be used in further processing. Each document in the Classification Training Batch is marked as either "For training" or "For testing." The program will assign these statuses automatically, but you can always change them manually.

Once the program has accumulated 500 pages per class, it will stop accumulating documents for auto-learning (otherwise, auto-learning would never cease). If a new class is added, documents of this class will be added to the Classification Training Batch and the classification model will learn to detect the new class.

Besides improving classification models, the auto-learning feature makes it easier to add new document classes to classification models. To enable the verification operator to add new classes, the project administrator must first add a Document Definition for the new type of document. Now the verification operator will be able to assign the new class to documents by modifying the Document Definition. Once the verification process is completed, auto-learning will start automatically.

To enable the auto-learning feature, open the Project Setup Station, click **Project > Project Properties > Workflow** and select **Allow training**. Finally, right-click the **Classification Training Batch** and select **Unlock Training for Operators** on the shortcut menu (this is necessary because once a classification model has been trained by the administrator it is locked to prevent training by operators).

2.3 Combining classification models for hierarchical classification

When you deal with a hierarchy of classes, you may combine several classification models. For example, the 1^{st} level of your hierarchy can differentiate between contracts, receipts, and invoices; one of the nodes at the 2^{nd} level of your hierarchy can sort contracts by subject; and maybe a 3^{rd} level can identify contract variations to suggest appropriate models for data extraction.

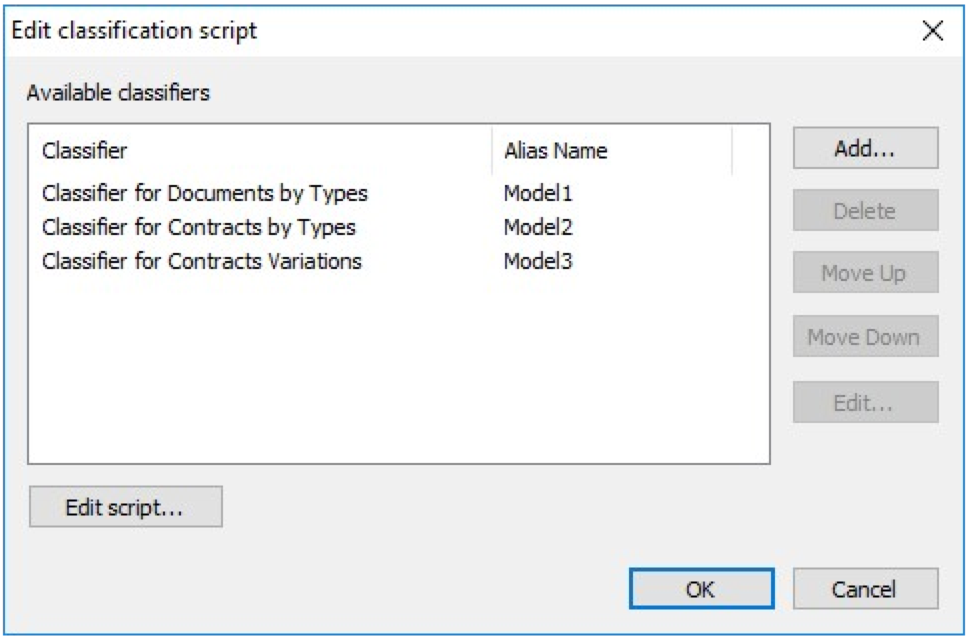
Having independent classifiers in the nodes of the cascade allows you to:

- Simplify the training setup (when you train each model you deal only with limited number of classes).

- Have different classification priorities for different classification tasks. If you want high precision in document type identification (for example a passport), but still want to identify the first page of the document inside a document type with high recall.).

- Speed up training (this is important, because when adjusting your classification solution to real-life documents you may have to retrain it again and again, and if you have several classification models, you can re-train only the one that produces an error).

You can combine several classification models using a classification script. First of all, you need to declare, in the script properties, all the classification models that you are going to use (they are represented by classification training batches). Use aliases so that your script is not affected by a changed classification model name.



The script provides a **IClassifiers Classifiers** object with two methods:

bool Has (string alias) – Checks if there is a classification model with the specified alias available for the script.

IClassifier Get (string alias) – Returns the classification model with the specified alias.

The **IClassifier** interface has a **ClassifyPage (IPage page)** method that allows you to classify a page using all the classification models declared in the script properties.

```
string result = "";

If (Classifiers.Has("Model1")) {

    IBatchTypeClassifierResult model1_result = Classifiers.Get("Model1").Classify(Page);

    If (model1_result.ClassName != "")  result = model1_result;

    If (model1_result.ClassName == "Contract") {

            IBatchTypeClassifierResult model2_result = Classifiers.Get("Model2").Classify(Page);

            If (model2_result.ClassName != "")  result = model2_result;

    }

}

Page.ResultClassName = result;
```

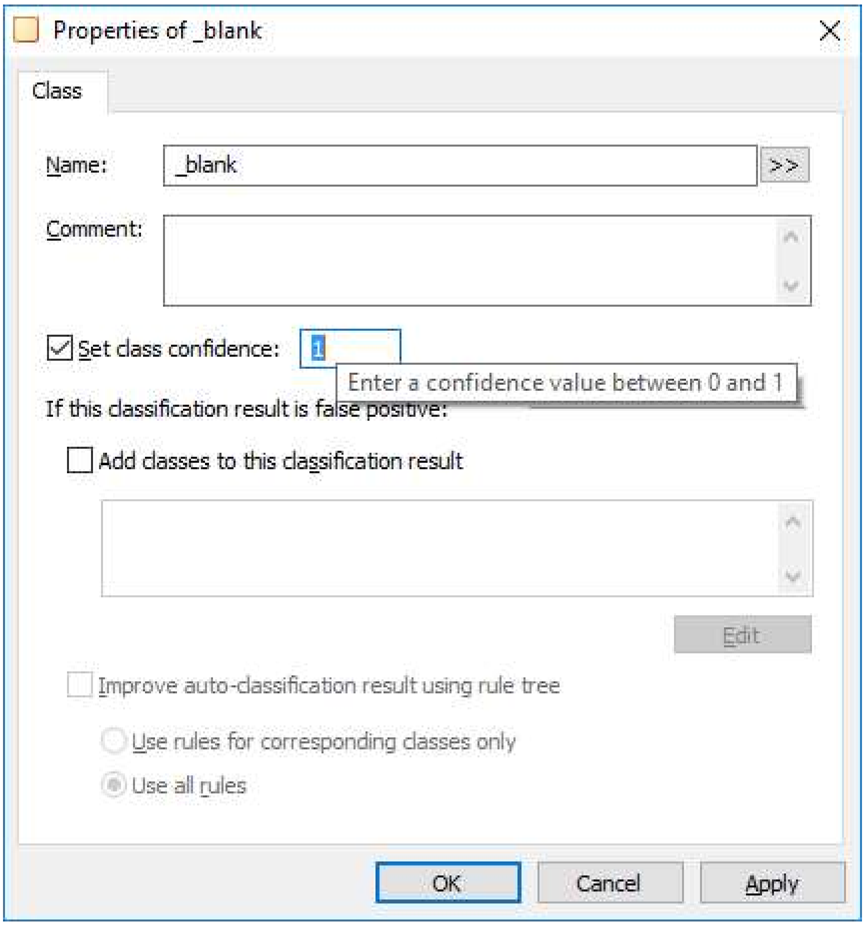
Note that when you combine several models in a classification script, autolearning cannot be used.

2.4 Handling "garbage"

Example 1. To train a classification model, you need to have at least two classes. But what if you just need to detect, say, passports among other different types of documents? In this case, you need to train a "Passports" class and an "All Other Documents" class. Be sure to provide a training set that is representative of the real-life stream of incoming documents.

Example 2. If you need to sort documents, say, into invoices and receipts, but may also receive some unknown document types that your business process doesn't expect, it is best to create an "Unknown" class for such unexpected documents. This will improve classification quality.

Example 3. Let's assume now that in Example 2 you train your model for high precision, because you don't want to misclassify an invoice as a receipt, as the cost of error is very high. At the same time, it is much more preferable to have some unknown document classified as an invoice than to overlook an actual invoice. To sway the classification model in favor of the "Invoice" and "Receipt" classes, you can specify a confidence level smaller than 1 for the "Unknown" class. The corresponding option can be found in the properties of the class.



As confidence levels smaller than 1 mean that the classification model is more likely to decide in favor of some other class, you may also wish to specify a confidence level for the “All Other Documents” class in Example 1 above.

2.5 Sharing a classification model between batch types

Starting from this release, batch types no longer “own” classification models, which are now stored inside classification training batches. You can use the same classification model across multiple batch types if it is needed for your business process. Select one classification model for each batch type. If auto-learning is enabled, a shared model will be improved from processing documents within all the batch types where it is used.

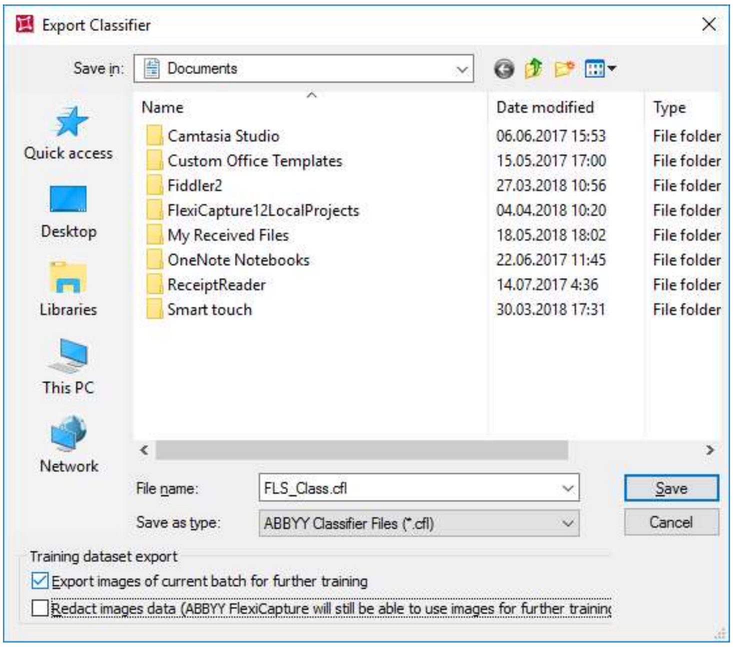
2.6 Reusing training results in other projects

You can use the same classification model in multiple projects. For example, you can use a domestic dataset, probably collected from different customers, to train a model and then deliver this model to a specific customer.

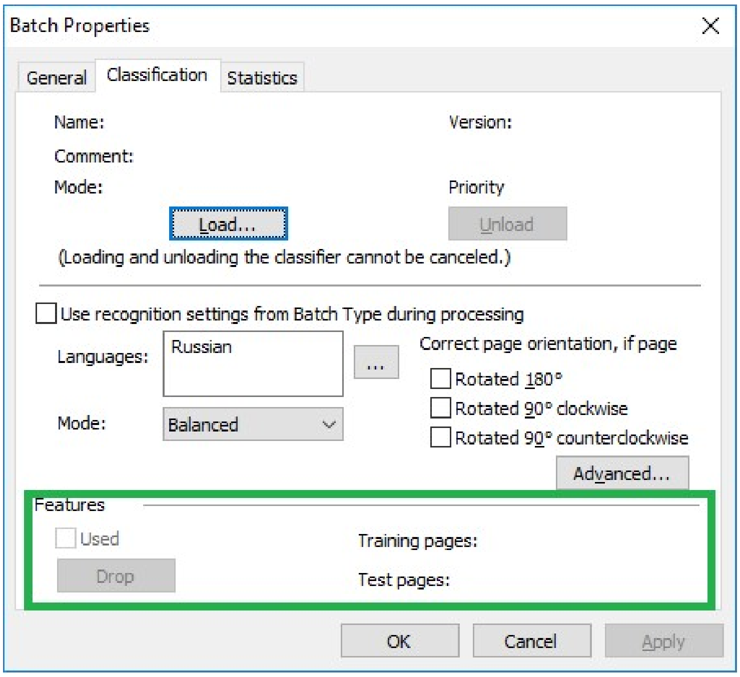
However, in order to continue training the model for his own documents, the customer must also have access to your domestic dataset, otherwise he will be training the model from the scratch and may initially get low classification quality.

FlexiLayout Studio allows you to export a classification model alone or with its domestic dataset. Similarly, FlexiCapture allows importing a classification model alone or with its domestic dataset.

Before providing a dataset that was collected from different customers to someone else, you need to anonymize it. An anonymized dataset does not contain any documents that a human can read. Instead, it only contains features calculated for classification purposes. Form the point of view of the technology it is no different from a non-anonymized dataset, but its documents cannot be read by humans. To export an anonymized dataset from FlexiLayout Studio, enable the “Redact image data” option



FlexiCapture allows importing an anonymized dataset together with its classification model. The anonymized dataset will not be available as a list of documents in the batch. Instead, only the size of the dataset will be shown in the batch properties.



The anonymized dataset can be switched on and off to estimate its contribution to the quality of the model. Or it can be discarded altogether if the documents collected by the customer provide good classification quality.

Note that anonymized datasets may become incompatible with future major updates to FlexiCapture. For this reason we recommend discarding such datasets if enough sample documents have been collected on the customer side.

3. Facilitation of the reverse matching scenario

3.1. Sharing registration parameters between documents and FlexiLayouts

You can pass a document's registration parameters to a FlexiLayout. It is possible to pass registration parameters from one FlexiLayout to another if both are used to process the same document in the same Document Definition. For this purpose, input and output parameters are now supported in FlexiLayouts.

This feature may be useful in the following scenarios:

- Reverse data matching allows a supervisor to verify that extracted data exists in the scanned documents after manual or other kind of processing. FlexiCapture compares data from the database with data on the document images by transferring document input parameters to FlexiLayout.
- Search information about specified subject (person, company) in the document. It is now possible to transfer information with is expected to be available on the document into FlexiLayout and verify if this information can be really found.

To use this feature, complete the following steps:

1. In the FlexiLayout to which additional information must be passed, create Input Named Parameters.
   To do this, right-click the FlexiLayout in ABBYY FlexiLayout Studio, select **Properties** on the shortcut menu, click the **Input Named Parameters** tab, and then click **Add New**.
2. In ABBYY FlexiCapture, for the document to be matched with the FlexiLayout, specify registration parameters with names identical to the names of the Input Named Parameters created in the FlexiLayout. To do this, right-click the document, select **Properties** on the shortcut menu, click the **Registration Parameters** tab and then click **Add**. In the dialog box that opens, enter the name and specify "String" as the type.

Specify the values of the registration parameters. This can be done either manually or automatically. To specify the values manually, open the document properties, click the **Registration Parameters** tab, then click the field in the **Value** column of a registration parameter and enter a value.

Values can also be specified by using a script at the automatic stage or during import.

4. Classification based on a database of companies

Classification based on companies is now available in all projects.

This feature allows you to associate document with some company from the Data Set, for example, we can classify bank statements by banks having a database of banks. Normally, document issued by different companies have different layouts, in this case such classification will allow using training of field extraction by layout: documents from each company will be trained independently. List of companies can be imported from external database or collected during data entry as operators can add and edit records during verification. FlexiCapture easily supports Data Sets with more than 100 000 companies.

To be able to use this feature, you need to enable the corresponding option in the section properties and add the available company details to the section **Variants** data set:

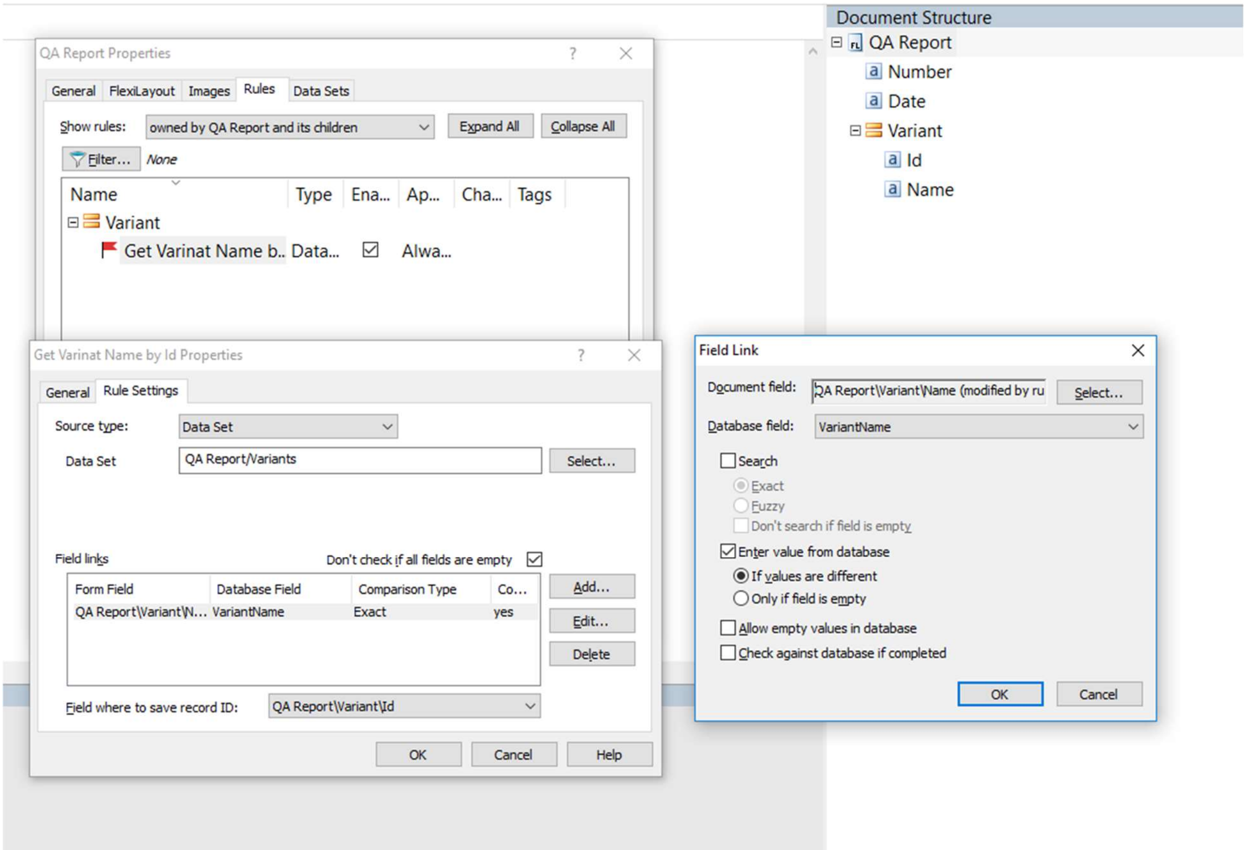
1. Open the Document Definition editor, right-click the desired section, and select **Properties** on the shortcut menu.
2. On the **Dataset** tab, click the **Variants** data set.
3. Select the **Use database of companies** option to add fields to the data set.
4. Establish a connection to the database and map the fields.
5. If a field is not mapped to any field in the database and does not contain any data, it will not be ignored by the classifier.
6. This feature will only work if the variants data set contains at least one populated field (e.g. a company name).

Classification based on a database of companies does not require additional training. The program will look for a company after its details are added to the variants dictionary. For this reason, you may wish to allow your operators to add new companies or edit existing company details during verification.

To do this, you need to configure a look-up rule for the variants dictionary of the section:

1. In the Document Definition editor, create a Variant group and then create two fields in this group—a service **Id** field and a text **Name** field.
   For the **Id** field, specify **Flexible Section Variant ID** as the source.
   To identify a company, the program will use its ID, but it is more convenient to display the company name to the operator, because usually it's the company name that the operator will see on the image.
   You can use any other fields instead of the **Name** field.
2. At the level of the Variant group, create a database check rule.
   Right-click the groups and select **Properties** on the shortcut menu. On the **Rules** tab, click **New Rule** and select **Database Check**

3. Select **Dataset Variants** as the source. Specify the **Id** field as the **Record ID.**
   Click **Add** and specify **Name** as the document field and **VariantName** as the database field. Disable the **Search** option and select **Enter value from database** and **If values are different**.



4. Add a button to your form.
   Right-click where you want to add your button on the form and select **Insert > Button**.
   On the **General** tab, enter a name for your button.
   On the **Format** tab, select your previously created database check rule.

5. Enterprise-ready platform

5.1. Oracle support

You can create a new Oracle database or connect to an existing one on the Administration and Monitoring Station. To do this, click **Settings > Application Server,** select Oracle for the database type, enter the desired alias and the server instance name, and fill in the authentication details.

Select the **Use external file storage** option and specify the path to the file store. For details, see http://help.abbyy.com/en-us/flexicapture/12/monitoring_console/mmc_appserv_storage.

Use the **Test Connection** button to check the entered values. If the test is successful, click **OK**.

Note: Be sure to use the Database Server Authentication as Windows authentication is not supported.

5.2. Microsoft SQL Azure support

You can connect to an Azure database on the Administration and Monitoring Station. To do this, click **Settings > Application Server,** select Azure for the database type, enter the desired server instance and database name, and fill in the authentication details.

Select the **Use external file storage** option and specify the path to the file store. For details, see http://help.abbyy.com/en-us/flexicapture/12/monitoring_console/mmc_appserv_storage.

Use the **Test Connection** button to check the entered values. If the test is successful, click **OK**.

6. Connectors

6.1 Export to SharePoint

The SharePoint connector has been rebuilt using the latest API from Microsoft. ABBYY FlexiCapture 12 natively supports SharePoint Online (with all the possible authentication methods), SharePoint 2016, and SharePoint 2013.

ABBYY FlexiCapture 12 does not support SharePoint 2010 and below.

6.2 Universal export connectors

To facilitate the creation of export connectors to backend systems, a universal programming interface has been introduced. If a partner implements this interface in his .Net code and adds his DLL files as attachments to a Document Definition, then a new type of export profile will become available in ABBYY FlexiCapture and the end-user will be able to set up export settings via the GUI.

6.2 New connectors

Connectors to the following systems are now available: Laserfiche, Appian, M-Files, Pega, UiPath, and BluePrism. They are distributed separately from the ABBYY FlexiCapture 12 distribution package.

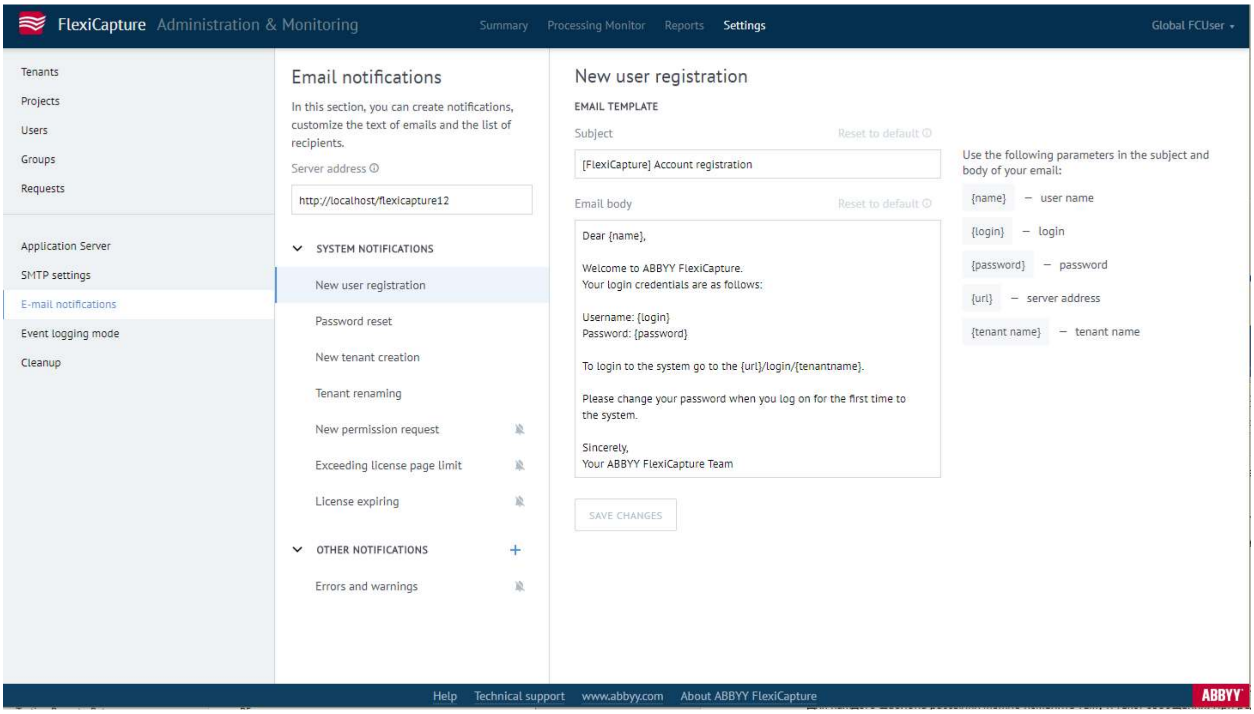
7. E-mail notifications

7.1. Improvements

You can configure e-mail notifications on the Administration and Monitoring Station (**Settings -> E-mail Notifications**), where the following notifications are available:

- New user registered
- Password reset
- New tenant created
- Tenant renamed
- Permissions requested
- Running out of pages
- License about to expire
- Errors and warnings *(New)*

Now all e-mail notifications can be edited in the GUI. You can view the template for each type of e-mail notification and modify it if necessary. The subject and e-mail text can be changed for each notification template. Special tags can be used in the text which the system will automatically substitute with appropriate values. You can always revert to the default text.



For “Permissions requested,” “Running out of pages,” “License about to expire,” and “Errors and warnings” notifications, both users and user groups can be added to the list of subscribers. The “Send to” field also works as a search field, allowing you to find a user and add him/her to the list of subscribers.

“Permissions requested,” “Running out of pages,” “License about to expire,” and “Errors and warnings” notifications can be turned on or off by selecting the **Send e-mails** option.

7.2. New “Errors and warnings” notification

“Errors and warnings” is a new type of e-mail notification added in this release. It can be configured to report errors and warnings in one or several projects. You can add users and user groups to the list of subscribers, specify how often the notification should be sent, and set the number of errors to be displayed in the e-mail message.

You can create several templates of this type of e-mail notification for different projects and users.

7.3. E-mail notifications now available in tenants

E-mail notifications are now available in tenants. The administrator of a tenant can configure the following e-mail notifications in the tenant: “New user registered,” “Password reset,” “Permissions requested,” “Errors and warnings.”

“Running out of pages” and “License about to expire” notifications currently cannot be configured in tenants.

E-mail notifications are configured in tenants in the same manner as in the default tenant.

8. Reports improved

8.1. Batch details in the General Operators Report

The General Operators Report now includes a new option, **Batch details**, that allows you to add additional information to the report.

Selecting the new **Batch details** option will add **Batch name** and **Page count** columns to the General Operators Report. The **Batch name** column displays the name of the batch that the operator worked with. The **Page count** column displays the total number of pages in the batch. By default, the **Batch details** option is not selected.

A new column named **Work time** has been added to the General Operators Report. It displays the time spent by the operator working on a particular stage. For scanning and rescanning stages, a dash will be displayed in this column.

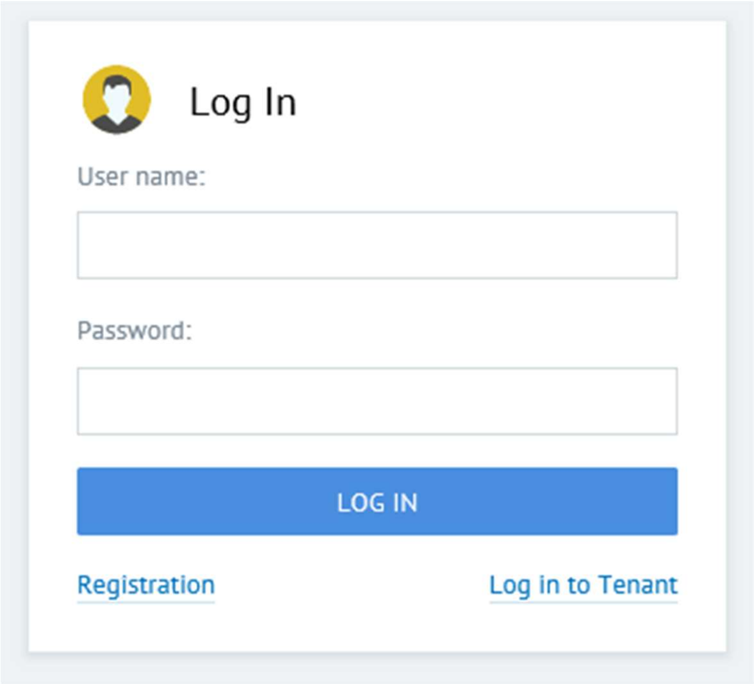
For site Performance, Processing, and General Operators Reports, you can now gather statistics about scanning, rescanning, and custom stages.

8.2. Report templates

Now for each type of report you can create your own set of templates with different report settings. For example, two separate templates can be created based on the General Operators Report: one with statistics about the verification operators and one with statistics about the scanning operators. These templates can be used to generate daily reports to evaluate the performance of each verification or scanning operator. Report templates are saved for each operator individually.

9. Logging in to the multi-tenant system

To log in to the system as a user of a particular tenant, you must click the “Log in to Tenant” link on the **Log In** page.



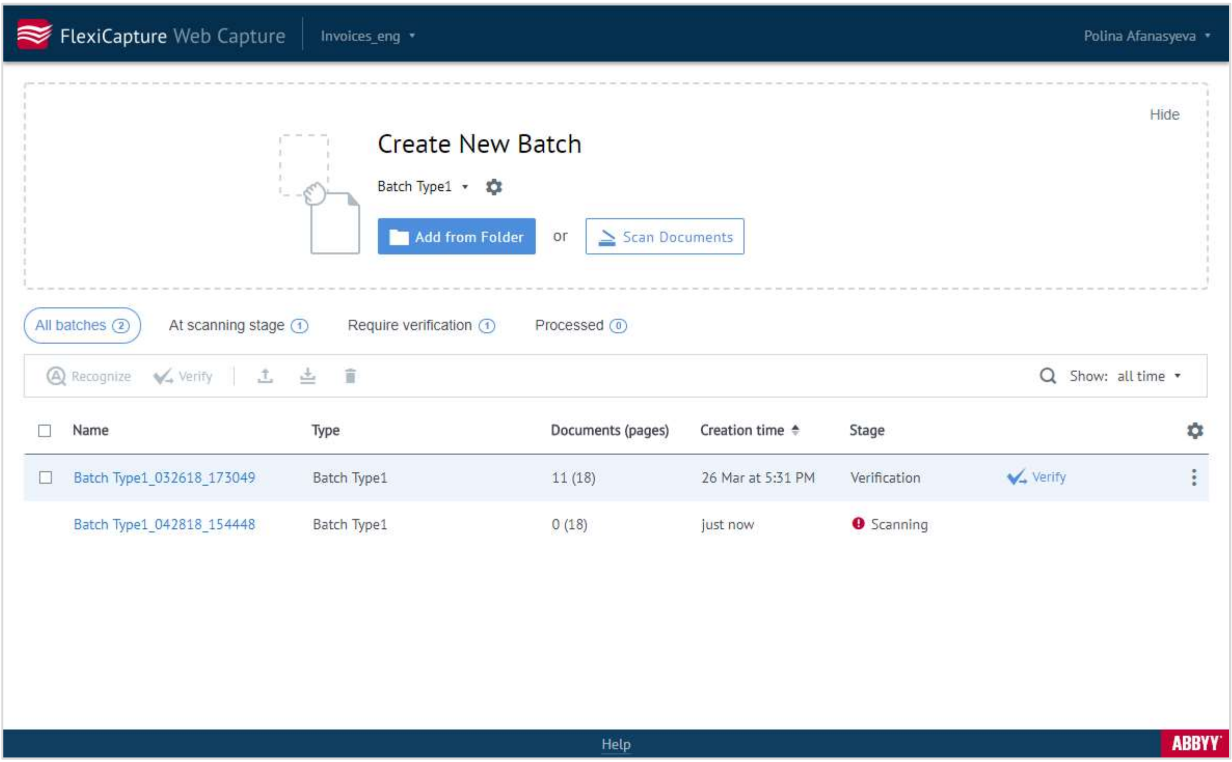
On the **Log in to Tenant** page, enter the name of the tenant in the corresponding field, and the system will redirect you to the tenant login page, where you must specify your user name and password.

10. HTML5 web stations

10.1. Web Capture Station

The Web Capture Station is now available in HTML5.

The interface of the Web Capture Station allows operators to scan documents or add documents from a folder, submit input data for recognition, verify and correct recognized data, export verified documents to an external system, and load results for review. Each operator has access only to his/her own batches.



To start working on a batch, the Web Capture Operator should select a batch type, add images from a folder or scan documents, specifying scanning settings if required.

Once the images have been added, the operator will be redirected to the scanning page. On the scanning page, the operator can check the added images, create documents, edit (e.g. rotate, flip, convert, invert, despeckle, deskew, crop or redact) images, and fill in the registration parameters for the batch and documents if required. Next, the operator can send the batch to the server for further processing.

The processed batch will appear in the list of batches on the Web Verification Station. The operator can verify the batch by clicking Verify. Next, the operator will be redirected to the verification page.

On the verification page, the operator can check that the documents have been assembled correctly and the Document Definition has been properly matched, verify the extracted data, and correct any recognition and rule errors.

The verified batch can then be exported.

The Web Capture Station allows downloading export results. To download export results as a ZIP archive, select the required batch from the list and click .

10.2. Web Verification Station

10.2.1. Custom data form

A custom data form is now available on the Web Verification Station and includes all the settings configured in the Document Definition. The following form elements and parameters can be modified in the Document Definition editor:

- position and size of fields
- font size for fields and labels
- static text
- table and table settings
- ability to hide/show empty columns in the table
- ability to add an instance of a repeating group/field anywhere on the form
- custom tab order
- custom buttons

10.2.2. *Invoice line items*

Invoice line items are displayed as a table on the data form. A line item may include such fields as order number, article number, description, quantity, unit price, total price, etc. By default, only those columns are displayed on the data form whose fields have been detected on the image or whose display has been turned on in the Document Definition by the administrator.

If invoice line items have not been captured automatically or have been captured incorrectly, the operator can capture them manually or try fine-tuning the automatic capture feature.

To mark up a line item, the operator must draw a rectangle around the line item on the image and then select **Line Items** on the shortcut menu. Holding down the left mouse button or using the **Add Separator** tool, the operator can draw vertical and horizontal separators. A plus sign will appear at the top of the column. When the operator clicks the plus sign and selects the name of the column from the list, the corresponding fields will be automatically created inside the column.

Then the operator creates regions for the fields in the first line item in the column. When the operator starts creating regions for the next line items in this column, the program will display field prompts based on the information it has accumulated about the preceding line items.

When the first line item is marked, the operator can use the **Continue Line Items** command for more reliable detection. The program will use information about the position of the first line item and about the positions of the field regions within the line item.

If the result is unsatisfactory, the operator should correct the position of the first incorrectly detected line item, specify the fields of this line item, and apply the **Continue Line Items** command once again.

This will cause the program to detect the remaining line items automatically. For more reliable detection, the program will use information about the position of the current line item and about the positions of all of the preceding line items.

If the result is still unsatisfactory, the operator should manually draw a rectangle enclosing all of the line items and then draw horizontal separators to separate the individual line items.

If the line items have been detected correctly but some of their fields have not been properly identified, the operator can use the **Reanalyze Line Items Below** command, which is available on the shortcut menu of the line item's region.

This command takes into account the positions of the fields in the current line item and the line items that precede it, which are deemed to be provided (or adjusted) by the operator. Using this information, the program identifies the fields in the line items below the current line.

If the line items have been detected correctly but some of their fields have not been properly identified, the operator should arrange the field regions in the first line item and select the **Reanalyze Line Items Below** command.

The program will automatically detect the fields in the remaining line items, without changing the positions of the line items themselves. For more reliable detection, the program will use information about the positions of the field regions in the first line item.

If the result is still unsatisfactory, the operator should adjust the position of the fields in the first line item where the fields are detected incorrectly and select the **Reanalyze Line Items Below** command for this line item.

The program will automatically detect the fields in the remaining line items, without changing the positions of the line items themselves. For more reliable detection, the program will use information about the positions of the field regions in the current line item and in all of the preceding line items.

10.2.3. *Working with complex regions*

When verifying document data, the operator can work with complex regions and continued regions.

A complex region is a region displayed as a set of multiple words located on multiple lines. A continued region is a region whose values are located in various places on a page or on multiple pages.

The operator can create a complex region on the image and bind it to a desired field, modify, or move it.

10.2.3. Database lookup for all fields

A database check rule compares a captured field value against the values stored in a database. If a captured field value is found among the database values, the check is passed. Otherwise, the rule returns an error.

You can use this rule to check if there is an entry for a captured value in your database, to carry out fuzzy searches in your database even if there are typos or errors in fields, to replace a field value with its corresponding record from the database, and to copy certain values from the database into document fields.

Now a database check rule can be implemented for all fields on the data form, including Business Units, Vendors, and Purchase Order.

For instructions on creating a database check rule, please see this section of the online manual: http://help.abbyy.com/en-us/flexicapture/12/distributed_administrator/rule_bd.

A new database lookup window allows you to search for a value, view the data in the data set, and select data. The operator can edit the data set or add a new one (if allowed by the administrator in the project settings).

10.2.4. Keyboard shortcuts

Keyboard shortcuts are now available on the Web Verification Station. These keyboard shortcuts are currently hardcoded and cannot be edited.

Queues View	
Get Task	Ctrl + G
Batch Editor	
Apply Changes	Ctrl + Q
Expand All/Collapse All	Ctrl + M + O/Ctrl + M + P
Create Document	Ctrl + Alt + N
Merge Documents	Ctrl + Alt + M
Clear Analyzing Results	Alt + Del
Rotate 90 Clockwise	Ctrl + >
Rotate 90 Counterclockwise	Ctrl + <
Image Editor	
Fit to Width	Ctrl + 2
Fit to Height	Ctrl + 3
Actual Size	Ctrl + 1
Zoom In	Ctrl + mouse / Ctrl + Num+
Zoom Out	Ctrl + mouse / Ctrl + Num-
Working with Tables	
Draw Vertical Separator	Ctrl + \
Draw Horizontal Separator	Shift + \
Delete Separator	Alt + \
Move Active Region	F2 + up/down/right/left arrow
Delete Active Region	F2 + Del
Data Form	
Check Rules	Ctrl + Shift + S
Open Drop-Down List	Down arrow
Open Suggestions	Ctrl + down arrow
Go to Beginning/End of Document	Home / End
Next Item to Verify	F4
Previous Item to Verify	Shift + F4
Next Low-Confidence Character	F8
Previous Low-Confidence Character	Shift + F8
Next Error	F6

Previous Error	Shift + F6
Error Filter Switching	Alt + 1,2…
Copy	Ctrl + C
Paste	Ctrl + V
Delete	Del
Next Document with Errors	F9
Previous Document with Errors	Shift + F9
Next Document	Ctrl + D
Previous Document	Ctrl + Shift + D
Go to Beginning/End of Page	Page Up/Down
Go to Beginning of Document	Home
Go to End of Document	End
Confirm Field	Enter
Operations on Tasks	
Complete Task	Ctrl + L
Reject Task	Ctrl + J
Postpone Task	Ctrl + Alt + L
Send to Stage	Ctrl + Alt + Q
Send to Exceptions	Ctrl + Shift + X
Rescan	Ctrl + Alt + P
Recognize	Ctrl + R

10.2.5. Changes in the API related to the integration of the Web Verification Station into third-party systems

In the integration URL of the Web Verification Station, there is now a **mode=mini** parameter that you can use to switch off the header and footer of the Web Verification Station.

To load the Verification Station GUI and functionality into the <iframe> element, set the value of the **src** attribute to:

*<stationUrl>/WinAuth/Verify?projectId=[projectidentifier]&roleId=[roleidentifier]&stageId=[stageidentifier]&taskId=[task identifier]&returnTo=DeadEnd?**mode=mini***

For more details about integrating the Web Verification Station into third-party systems, please see this section of the online manual: http://help.abbyy.com/en-us/flexicapture/12/developer/verification_integration.

10.2.6. "Autoreceive" mode for tasks

To continuously process document batches, operators can use the "autoreceive" mode for tasks. This mode is enabled by selecting **Automatic Task receive mode** in the task view.

The operator can activate automatic receipt of tasks to process batches either from a general queue or from a specific queue.

- If a queue contains tasks that can be accessed by the operator, he/she will receive them after the current task is completed.
- If the first task is received from a general queue by clicking **Get any task**, with the autoreceive mode enabled, the next task will be taken from the general queue.
- If the first task is received from a specific queue, with the autoreceive mode enabled, the next task will be taken from that queue.

If a queue is empty after the operator closes, sends or postpones a task, a standby mode is activated. In this mode, the operator will see a standby screen. If a new task appears in the queue, its name will be displayed in the title of the browser tab.

If the operator is in the queue view and has no tasks to process at the moment, the operator may activate the standby mode by clicking **Wait for task**.

If the current task is canceled, the Automatic Task receive mode will be reset.

10.2.7. *Database lookup using a custom script*

Before you can configure a custom script to search the database, you must complete the following steps:

- Open the Project Settings Station and the open the Document Definition editor by clicking **Project → Document Definitions...**
- Click **Document Definition → Document Definition Properties** and on the **Datasets** tab, click the **Add...** button.
- Create a dataset, enter a few records, and click **Close**.

Your custom script will be added to the Scripts Editor for Web.

A template script for database lookup can be found at http://<ApplicationServer>/FlexiCapture12/Verification/Scripts/CustomScripts/minilookup.js, where <ApplicationServer > is the name of the computer on which the Application Server is installed.

You can create your custom script from the template script by replacing the values in the following lines:

var dbCacheName = 'DataSet1', where DataSet1 is the name of the previously created dataset, and

var fieldNameInCache = 'Field2', where Field2 is the name of the dataset field from which the value will be copied.

10.3. Web Scanning Station

10.3.1. *Global scanning settings*

The administrator can specify settings for each batch type and save them on the server. Then all the scanning operators will use the same settings for each batch. This will help avoid mistakes, and operators will not need to set up batch types and their settings manually.

10.3.2. *Integration into third-party systems*

The Web Scanning Station can be integrated into a third-party system by means of single sign-on technology (using SAML 2.0 and Windows authentication). Users will interact with the integrated Web Scanning Station in the same manner as they interact with a non-integrated station. The URL parameters are customizable and support headless/footless mode.

To load the Web Scanning Station into the <iframe> element, set the value of the **src** attribute to:

<stationUrl>/Manage/winAuth/OpenOrCreate?projectId=[Id_of_project]&batchTypeName=[name_of_batchtype]&batchName=[name_of_batch]&regparam=[parameter:value], where:

Manage – This method turns off the header and the footer on the Web Scanning Station.

winAuth – Windows authentication parameter.

OpenOrCreate – This command opens an existing batch that has not yet been sent or creates a new batch.

projectId – The ID of the project (obligatory parameter).

batchTypeName – The name of the batch type (obligatory parameter).

batchName – The name of the batch (optional parameter). If this parameter is not specified, the batch will be named according to the batch naming scheme specified in the batch type settings on the Web Scanning Station.

regparam – Registration parameter (optional). This parameter can be repeated (if several registration parameters should be filled in). It has the format “name:value” and should be specified in the batch type settings of the Web Scanning Station. If no value is specified in the settings of the Web Scanning Station, this parameter is ignored when creating a batch.

To get a task for re-scanning, use the following URL:

<stationUrl>/Manage/winAuth/GetRescan?projectId=[Id_of_project]&batchName=[name_of_batch]&regparam=[parameter:value], where GetRescan gets a task for re-scanning.

After closing or sending the batch, postMessage will pass an FC_Scanning_TaskClosed message to the parent element.

URL encoding should be used to pass non-English characters as parameter values in the URL (e.g. Japanese characters).

10.3.3. Image enhancement operations

The Crop and Redact operations are now available for editing images. The Crop and Redact operations are located in the image editor and can be applied to one or several documents.

10.4. TLS1.2 support

Web stations now support the TLS 1.2. protocol. To use the protocol with the HTTPS connection, follow the steps described below:

- Enable TLS 1.2 using IIS Crypto.
- Enable the protocol on the server side for the Application Server.
- Establish client-side protocols for other computers where ABBYY FlexiCapture components are installed.
- Disable any other protocols if necessary.

Important! Restart your computer after applying the changes.

The tool changes register keys in the following branch:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\SecurityProviders\SCHANNEL\Protocols

10.5. SSL and HTTPOnly flags

HTML5 web applications now support SSL and HTTPonly flags if they are specified in web.config. Now the HTTPOnly flag is turned on for all cookies except FCUserLang regardless of the settings. The SecureSSL flag can be turned on in web.config as follows: <httpCookies requireSSL="true" />.

11. Customizing HTML5 web stations

11.1. What can be customized

By customizing HTML5 web stations you can change the appearance of web pages, some UI elements, and the way users switch between stations.

A customization page is a page with an iFrame element containing a web station.

Only the workspace of a web station is opened in iFrame, without its header or footer. The layout of the customization page can be changed to meet the requirements of your company brand book. The customization page can include action buttons that can manage the content of the iFrame, web links, logos, etc.

Possible changes:

- Change logo, header, and footer
- Change links to company web site and other resources
- Change station icon in browser
- Change page name in browser
- Hide language drop-down list
- Select necessary role automatically on Web Verification Station
- Preserve login data when switching between web stations
- Change logoff period

11.2. Enabling customization

The customization page can be accessed at the following URL:

http(s)://<server>/FlexiCapture12/<station>/WinAuth/Custom, where

<server> is the computer name with the web server client,

<station> is the name of the web Scanning or Verification Station, and

WinAuth is the Windows authentication parameter.

11.3. Files required for customization

The layout of the customization page can be configured by modifying the following file:

C:\inetpub\wwwroot\Flexicapture12\<station>\Views\Custom\Index.cshtml

Other required content files should be loaded into:

C:\inetpub\wwwroot\Flexicapture12\<station>\CustomContent

JS and CSS files will be automatically loaded into the custom page.

11.4. Possible changes

Change logo, header, footer, page name, page icon, add web links, and hide languages

To change the logo, header, and footer, to add links to your company web site and other resources, to change the station icon or page name as displayed in browsers, and to hide the language drop-down list, you must change the corresponding values/strings in the **Index.cshtml** file located at:

C:\inetpub\wwwroot\Flexicapture12\<station>\Views\Custom\Index.cshtml

This file includes detailed comments.

Select necessary role automatically and preserve login data when switching between stations

Using the server API (http://help.abbyy.com/en-us/flexicapture/12/developer/fcintegration) you can customize:

- Switching between web stations using action buttons
- Automatic selection of the required project and role
- Preserving user login and saving session data

Change logoff period

To specify a session duration to log off inactive users, you must change the value of the **SessionInactiveLifetimeInMinutes** parameter in the **web.config** of the corresponding web station.

If a user is inactive during the time period specified in the **SessionInactiveLifetimeInMinutes** parameter, the user will be logged off and redirected to the login form.

Other possibilities

You can set up a web station to load into iFrame. The user can use the customized page of one web station to access other web stations.

As a custom page has a web station integrated into it, you can also use API post messages on this page as described in this section of the online manual: http://help.abbyy.com/en-us/flexicapture/12/developer/fcintegration.

11.5. Customization templates

In ***C:\inetpub\wwwroot\FlexCcapture12\<station>\CustomContent,*** you will find files with customization templates:

- **Custom.css** – a template for layouts
- **Custom.js** – a template for web page
- **Icon.ico** – a template for page icon
- **Logo.png** – a template for page logo

The file ***C:\inetpub\wwwroot\FlexiCapture12\<station>\Views\Custom\Index.cshtml*** contains a template for the layout of a custom page.

All files include detailed comments.

You can change or delete these files or add other files as may be required.

12. Document Definitions

12.1. Document Definitions without automatic field extraction

A new type of Document Definition is now supported to address the following scenarios:

- Some document sets contain documents that cannot be recognized, such as hand-written notes, certificates, letters of authority, or other unstructured documents. It is critical to classify such documents and check if they are present in the document set, but there is no need to extract their fields.
- All incoming documents should be classified and saved to their appropriate storage locations, but data must be extracted from just a few of the document types.

In both of the above scenarios, document types will be assigned by the classifier or by the operator. Data will not be extracted, but it will be possible to enter key fields manually or from the image and to create a searchable PDF document if necessary.

To create this type of Document Definition, the following conditions must be met:

- The Document Definition must have no anchors or identifiers. Otherwise, it will be applied as a fixed-form Document Definition.
- No FlexiLayout loading or training is allowed. Otherwise, Flexible Description will be matched.

Such a Document Definition will be excluded from analysis that precedes its application, as the type can be assigned by the classifier or by the operator.

Such a Document Definition will not require any recognition if there are no fields with the **Has no region** option selected. However, fields can be created to support manual or key-from-image indexing. To get text by click from image, create at least one field with the **Can have region** option selected. In this case the document pages will be recognized and a text layer for indexing will be created.

A document can also be recognized at the export stage if a searchable PDF document must be created.

Use the **Unstructured** option in the Document Definition assistant to create a Document Definition of this type. Note, however, that if you add anchors, allow training, or load a FlexiLayout, the type of the Document Definition will change. You can also obtain a Document Definition of this type from other Document Definitions by deleting all anchors, FlexiLayouts, and disabling training in them.

12.2. New Document Definition assistant

The interface for creating a new Document Definition has been changed to support the following typical scenarios:

- Create a Document Definition for fixed documents
- Create a Document Definition for flexible documents using FlexiLayout or training
- Create a Document Definition for unstructured documents (select this item to create a Document Definition without automatic field extraction)
- Create a document set from Document Definitions

13. Improvements to document sets

13.1. Working with document sets within batch types

If you need to process document sets only within a batch type, you can now add document sets without child documents. In this case, loose documents will not be allowed in the batch type and will be put into sets.

13.2. Collecting information from repeated documents in the summary form

Document sets allow you to place valuable information from child documents and display it in the summary form for quick review. If some of the documents can recur in the set, you can now create a table where each line will display information from one document and the number of lines will equal the number of recurring documents.

To create such a table, open the Document Definition editor for the selected document set, open the summary section, create a new filed with a link by clicking **Create new filed > Link to existing field**, and select <u>document</u> as a source. As a result, a new group will be created, which will recur if the source document occurs more than once in the set. Then add to this group links to those fields from the same document that should be displayed in the summary. Now this group will be displayed as a table in the data form.

13.3. Opening digitally born documents on the Scanning Station

If you scan a document set and have a digitally born report that has to be processed in the set, you can add such this report on the Scanning Station. For ABBYY FlexiCapture to be able to use the text layer in this document, it must be saved and sent to the server. This can be achieved by selecting **Send text layer of pages** in the sending options when sending documents from the Scanning Station to the ABBYY FlexiCapture server.

Note: This feature requires Microsoft Office, LibreOffice or the built-in module that converts digitally born documents to be installed on the same computer where the Scanning Station is installed. If you use click-once installation or the remote station installer, the built-in module will not be installed by default. If you need this module installed on your remote stations, install the Scanning Station component from the ABBYY FlexiCapture 12 distribution package.

14. Improvements to desktop operator stations

General

1. Input of Japanese characters using the common tools has been improved. We have also added special languages and data types for Japan.
2. The online manual is now available on all stations.

Verification

3. Changes in rejected and incorrectly closed tasks can now be saved.
4. Navigation between verification modes is now more straightforward: from the autoreceive mode, you can switch between working and training batches.
5. The **Clear Analysis Results** and **Update to Latest Version** commands can now be applied to multiple batches.

Scanning

6. The type of a batch and its and registration parameters can now be changed at the re-scanning stage if such changes are allowed in the batch type.
7. The drop-down lists on the Scanning Station that contain values for registration parameters can now be quickly filtered by typing the first few letters.
8. A batch can be checked at any time before it is sent to the server by clicking **Check Batch Integrity** or pressing F9. The following parameters can be checked:
    - Required registration parameters
    - Document types and number of documents (to validate document sets)
    - Image resolution and page size

15. Security issues

ABBYY FlexiCapture 12 has passed penetration testing and the following issues have been identified and fixed.

15.1. SQL injections vulnerabilities

1-2. Vulnerable parameter = mask. The parameter is passed to the search now, and there is no concatenation with the body request. An exception will occur at the server level and the attack will fail.

3-5, 7. Vulnerable parameter = sortOrder. The **sortOrder** parameter has only the values of ‘true’ and ‘false,’ and these values are now not passed directly to database queries.

6.Vulnerable parameter = filter. It is now impossible to pass a value of a non-existing column in the sort parameter. An exception will occur at the server level and the attack will fail.

7-19. Vulnerable parameter = order. The **Order** parameter has only the values of ‘true’ and ‘false,’ and these values are now not passed directly to database queries.

15.2. Broken access control

The /FlexiCapture12/Login/Server/SevaUserProfile POST request no longer uses the FlexiCaptureTmsSts2 parameter. In addition, its editing now does not affect the behavior of the web application.

15.3. Cross-site request forgery

For all web applications (i.e. Web Verification, Web Scanning, Web Capture, Administration and Monitoring, and Login) the ASP.NET MVC AntiForgeryToken mechanism (https://msdn.microsoft.com/en-us/library/dd470175(v=vs.118).aspx) has been applied.

In addition, validation of AntiForgeryToken has been added for POST methods that require it.

15.4. Weak password policy

The password policy has been changed to comply with the Microsoft rules for configuring password policies (https://technet.microsoft.com/en-us/library/dd277399.aspx).

Now a user's password should contain English upper- and lower-case letters and Arabic numerals or special symbols (! @ # $ % ^ & * ( ) _ - + : ; , . > < =").

15.5. No account lockout

We have created a counter for login attempts that is incremented with each unsuccessful login attempt. When the user runs out of login attempts, the system will temporarily lock the user's account.

15.6. New version of the Wibu driver

We have switched to Wibu 6.60 as a known vulnerability was fixed there.

16. Synchronizing AD groups from different domains

It is now possible to synchronize (manually or by external web request) users with Active Directory when multiple AD domains are used. Previously, users could only be synchronized with one AD domain. Users from different domains should be imported and synchronized alternately. Groups will not be overwritten.

How to synchronize groups from different domains:

- During the first import form AD, the path to the server is saved in the ABBYY FlexiCapture database in the domain.com\group1 format. Synchronization will be carried out as usual. If a group name changes, the name will be replaced by domain.com\group1NewName in the database.
- It is only possible to synchronize groups of one domain (i.e. the domain that is is indicated in the setting). To synchronize groups from other domains, you must change the setting consecutively and synchronize.

After the upgrade to the current version:

- The database must be updated to the latest version.
- For all objects already imported from AD, the same path that was used during the last import or synchronization is used.
- During the first synchronization (but not import), the path will be replaced by domain.com\group1.

Troubleshooting:

After the upgrade, some groups may stop synchronizing. This can be caused by the incorrect path saved during the upgrade or a changed server name (e.g. the URL was replaced by an IP or an alias). Solution: Try importing the groups once again (there is no need to delete any of them).

17. New method of synchronizing ABBYY FlexiCapture groups and users accounts with AD

The API offers a new method for synchronizing ABBYY FlexiCapture groups and user accounts with AD.

Use the following cURL:

curl --ntlm --user winLogin:winPassword http://urlToMonitoring/winauth/ActiveDirectory/SynchronizeGroups --data "IsWindowsAuthentication=true" –data "SearchGroups=true" – data "SearchGroupsUsers=true" --data "SearchSubGroups=false" , where

- **winLogin** – user domain name
- **winPassword** – domain password
- **urlToMonitoring** – web Monitoring address

Note: The user whose **winLogin** and **winPassword** are used in the cURL must have the role of Administrator role on the web Monitoring.

The following parameters are supported:

- **Server** – domain name of AD
- **IsWindowsAuthentication** – Windows authentication on the AD server
- **UserName** – user's name for connecting to AD (not applicable if IsWindowsAuthentication=true)
- **Password** – users' password (not applicable if IsWindowsAuthentication=true)
- **SearchGroups** – search groups
- **SearchGroupsUsers** – search users in groups
- **SearchSubGroups** – search groups in groups
- **Query** – search string for finding groups or users in AD

Parameters should be specified in the following format: **data "parameter_key=parameter_value"**.

18. Performance counters

Performance counters have been added to measure the performance of ABBYY FlexiCapture 12 using the standard Performance Monitor.

Processing Server counters

Counter	Description
Free cores	The number of idle cores in the system. Zero means the system is fully loaded and there are no free cores.
Active cores	The number of cores currently used by the system. The value decreases when some of the stations stop processing tasks.
Pending tasks	The number of tasks that the server has detected but has not yet distributed among the Processing Stations.
Running tasks	The number of tasks distributed among the Processing Stations. This number should be equal or slightly greater than the number of cores. If there are several times more running tasks than cores, the Processing Stations are overloaded.
Held licenses	The number of used licenses for cores. This counter is useful for licenses with core limitations (however, they are counted for all licenses).
Import tasks	The number of import tasks performed by the system within the last 5 minutes. This number should not vary significantly throughout the entire uptime of the system.
Import time	The average time spent by the system on performing import tasks within the last 5 minutes. When the processing is stable, the average time should not vary significantly.
Recognition tasks	The number of recognition tasks performed by the system within the last 5 minutes. This number should not vary significantly throughout the entire uptime of the system.
Recognition time	The average time spent by the system on performing recognition tasks within the last 5 minutes. When the processing is stable, the average recognition time should not vary significantly.
Export tasks	The number of export tasks performed by the system within the last 5 minutes. This number should not vary significantly throughout the entire uptime of the system.

Export time	The average time spent by the system on performing export tasks within the last 5 minutes. When the processing is stable, the average export time should not vary significantly.
Other tasks	The number of other tasks performed by the system within the last 5 minutes. This number should not vary significantly throughout the entire uptime of the system.
Time spent on other tasks	The average time spent by the system on performing other tasks within the last 5 minutes. When the processing is stable, the average time spent on other tasks should not vary significantly.

Update 1

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 1	1299/24	12.0.2.1356	20560	16.1.251.66	2018.10.05

New Functionality

1. Localization

ABBYY FlexiCapture 12 provides user interface on the following languages: English, Russian, German, French, Spanish, Korean, Japanese, and Czech.

In Japanese localization we have corrected the following: export files and folders names are not translated and stay in English.

2. Displaying pictures of the fields on the data form on the web verification

Configuring the Document Definition the administrator can switch on the “Add field picture” option in the properties of the field. This function allows to display an image below the field on the data form. This setting is passed through the xml to the Web Verification station now. As a result, operator can see pictures of the fields on the data form during documents verification on web stations.

3. Displaying license parameters for tenant administrator

When a tenant uses a personal license, the administrator of the tenant can see the following license parameters: license number, pages limit, expiration date on the monitoring dashboard.

When a tenant uses a main license, the license widget is hidden on the Monitoring dashboard.

Note: The license number is also displayed on the license widget for the Administrator of default tenant on the Monitoring dashboard.

4. License email notifications within the tenant

Two new email notification are available for configuration within the tenant: Exceeding license limit and License expiring.

E-mail notification settings are available for the Administrator of the tenant on the Monitoring and Administration station (Settings -> E-mail notifications). E-mail notifications are available for editing through the interface.

The Administrator of the tenant can view the template for these types of email notification and modify it if necessary. The subject and email text can be changed for each notification template. Special tags can be used in the text, which will be automatically substituted by the system to required system value. Reset to default allows to revert to the default text.

The subscribers list can include both users and users groups. Send to field works both as a search string and as an addition string, allowing to find the user and add it to the subscribers list.

Exceeding license limit and License expiring email notifications can be turned on/off by selecting check-box Send emails.

5. Test Connection for SMTP Server Settings

Test Connection button allows to check the availability of the email server during SMTP server settings.

6. Hiding the Automatic Task Receipt Mode for the Web Verification Integrated to 3rd party System

The Automatic task receipt mode is hidden by the system automatically when the Web Verification is opened in a 3rd party system, i.e. the integrated URL contains mode=mini parameter.

7. Displaying a Product Version on Web Stations

A product version is displayed in the About modal window, which is called by clicking on the ABBYY icon in the lower right corner. The product version is displayed as a server version and web version.

The About window is also include the End User License Agreement (EULA) and the list of authors.

8. Localization of the Data Form on Web Verification

Localization of the data form is configured on the Project Setup Station in Document Definition Editor, Document Definition > Localization. You shall specify all data form fields on the desired language and import localization file back into the project. Web Verification station receives all document definition settings, including available localization into other languages. When the user switches the system language, the data form displays data form elements on specified language taking them from the localization file.

9. Fields with fixed width on data form

For fields which can contain a lot of text (e.g. large text segments to verify) it is possible to fix the width on the data form. To do that switch on the **Fixed width** option in the properties of a field's box on the data form while creating such field in the Document Definition Editor.

10. PDF text layer is used via Web Services API

PDF Text layer is now used when images are sent for processing via Web Services API.

11. Authentication using WebSeal on Scanning Station

Authentication using WebSeal on the Scanning Station is now supported.

12. Security enhancements

12.1. Password storage mechanism is reinforced

The hashing algorithm of users' passwords has been changed from SHA1 to SHA512 in the database.

12.2. Protection against Clickjacking

The Monitoring and Administration station cannot opened in <iframe> now. Thereby, the Monitoring station is protected from clickjacking attacks.

Bug Fixes

Issue description
PDF text layer is now correctly matched after crop or split image. Improvements are also done in the mode when text layer is obtained from PDF file instead of OCR, in particular for PDF forms
A user was not imported while importing a group if a user and a group are in different domains.
Issue of editing record in Database lookup is fixed.
Web.config settings were reset to default after upgrade.
"Does not belong to set" and "Is not empty" conditions were not saved for Field Condition in the export settings.
Paragraph elements detection is corrected.
The same license cannot be assigned to different tenants

Update 2

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 2	1299/25	12.0.2.1363	21603	16.1.251.66	2018.10.19

Bug Fixes

Issue description
When an image was uploaded via API the PDF text layer was not used in spite of the correspondent setting
IPE uploading project to server (due to corrupted training batch) is fixed

Update 3

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 3	1299/26	12.0.2.1370	22855	16.1.251.68	2018.10.30

New functionality

1. Localization into Serbian

Desktop stations UI is translated into Serbian.

Update 4

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 4	1299/27	12.0.2.1387	25159	16.1.251.75	2018.11.22

New functionality

1. Localization into Polish and Portuguese (Brazilian)

UI of web and desktop stations is translated into Polish and Portuguese (Brazilian).

2. Country support in Invoice projects

In Invoice project we have added more keywords for existing countries and keywords for new countries: Denmark, Sweden, Norway, Portugal, Argentina, Chile, Colombia, Peru.

Bug Fixes

Issue description
The processing stopped if a tenant license was updated
API method AddFileToExportResults didn't work in some cases
Invoice LineItems always required verification due to uncertainly found region

Update 5

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 5	1299/28	12.0.2.1398	27056	16.1.251.75	2018.12.05

New functionality

1. Localization into Chinese and Hungarian

User interface is now translated into Chinese and Hungarian.

Known localization issues:

- Dialog box to setup connection to SharePoint for export remains in English
- Reports on Administration and Monitoring station, the buttons Find and Next on the upper pane remains in English

2. Possibility to group checkmarks on fields verification

To make the verification of checkmarks groups more comfortable, we have added a possibility to place checkmarks in a table view on fields verification. The user can limit the number of checkmarks in a row and for example make it equal to number of checkmarks in a row on the image. To do that use setting **Number of checkmarks in row** in checkmarks groups properties on the **Data** tab.

3. Separate role for documents assembly

Now it is possible to create a role with rights to correct assembly but without rights to add or delete pages. To create such role, it is possible to add permission **Assemble documents** but switch off a permission **Modify batch content**. In case **Modify batch content** is added, the permission **Assemble documents** will be always added as well, so the operator could perform all operations with pages and documents.

Roles and permissions can be specified on **Project** > **Project Properties**… > **User Roles** tab > **New** or **Edit** button.

Bug Fixes

Issue description
Fields conditions "Does not belong to set" and "Is not empty" in export settings could not be saved.
Error occurred trying to recognize pages when there was a limitation of cores in FlexiCapture 12 Standalone.
Image Source in page's properties was not correct in some scenarios.

Update 6

Technical Information

Release	Part #	Build #	Installer Build #	OCRT build #	Release date
Release 2 Update 6	1299/29	12.0.2.1420	29512	16.1.251.84	2018.12.28

New functionality

1. Localization of rules and error messages of the data form on the web verification station

Localization of the data form is configured on the Project Setup Station in Document Definition Editor, Document Definition > Localization. You shall specify all data form fields on the desired language and import localization file back into the project. Web Verification station receives all document definition settings, including available localization into other languages.

When the user switches the system language, the data form displays data form elements on specified language taking them from the localization file.

Rules error messages will be displayed on the user's language also. But as rules are checked during recognition, the messages generated on Processing Stations will be displayed on the language of Processing Station. You can setup a language in the batch type properties, in this case error messages generated by any processing station will be on the specified language. Localization file for this language should be available.

Known issue: the language setting is not translated into UI languages and remains in English.

2. Changing invoice status via custom action on the web verification station

Invoice status is assigned to invoice based on rule checks and/or manual checks performed by operator. Invoice status is displayed in the corresponding field on the data form. The status of invoice is determined automatically by validation rules. When errors have been corrected, rule checks are carried out again and a new status is automatically assigned to the invoice.

The status of invoice can be specified manually. To allow the program to export invoice that does not satisfy some of the rules, the operator can change the status and leave corresponding comment.

3. PO matching on the web verification station

The purchase order matching allows to match invoice and corresponding order. By default, matching invoices against purchase order is disabled. During project settings user can enable this functionality on the project setup station. The purchase order table can be opened by clicking on Details button in the corresponding section on the data form.

The purchase order table can be customized: operator can switch off unnecessary columns and re-size it. Using search field operator can find the required information in the PO table. For the convenient verification strings between invoice on the data form and PO table can be bound.

4. Displaying end of users' license agreement (EULA) to users on web applications

End of users' license agreement (EULA) can be uploaded by Administrator of the system on Monitoring and Administration station (Settings –> User Agreement). Displaying of user agreement is switched off by default. Administrator can upload a user agreement for one or several languages. The uploaded user agreement is one for all tenants (including default tenant).

If the administrator has enabled the option to display a user agreement for users on web applications, users will have to read the agreement and accept it when first log into the web application.

5. Edge 64-bit support for HTML5 web applications

HTML5 web applications: Web Scanning, Web Verification, Web Capture, Login page, Monitoring and Administration work correctly in Edge browser. Supported Edge versions: 42.17134 and 44.17763.

Known issue: verification station works slowly with the tables containing several dozen of lines.

6. Repeated recognition without changing the task on web verification station

Now in cases, when operator wants to change document definition or page section or drop the result of recognition he can continue working under the same task. Besides, repeated recognition now possible without changing a task.

7. Help files on Japanese.

The list of translated documentation:

- System Administrator's Guide
- Distributed Administrator's help (includes FlexiCapture for Invoices help)
- Data Verification Operator's help
- Verification Operator's help
- Scanning Station help

8. Enhanced search of batches and tasks

- Search batch by sub-string of batch name is now available. Use * to replace any prefix and type a sub-string.
- Tasks search is enabled in the Explore queue view (ctrl+F command).

9. Possibility to select certificate using Mutual SSL
When you enable Mutual SSL on Application Server it is now possible to select certificate on the Processing Stations using certificate thumbprint.

Use the following register key to select certificate:

HKEY_LOCAL_MACHINE\SOFTWARE\ABBYY\FlexiCapture\12.0\FlexiBr
ClientCertificateThumbprint

Stations should be run under the user which has such certificate.

10. FlexiCapture for Invoices. Mode of choosing training batches
You can choose how training batches are connected to vendors. By default VATID or IBAN are used to make correspondence between vendor and training batch. In case they are missing, then VendorID (or VendorID+BUID, if there is dependence of VendorID on BUID) are taken in consideration - this allows to share training results between different records about one real supplier (unique VATID or IBAN). Now you can disable this behavior and rely only on VendorID (or VendorID+BUID) to make connection between training batch and vendor - that maybe useful if different BUs represents different customers, or different vendor records in master data about the same supplier in fact mean different type of invoices (goods/services, bills, etc.) The behavior can be switched on the Invoice Features dialog box (Document Definition properties > Invoice Settings > Additional Fields and Features).

Known issue: the new invoice setting is not translated into all UI languages and remains in English.

11. FlexiCapture for Invoices. Switching between extrapolation modes
New extrapolation for invoice line items is disabled by default, use the following register key to enable:

[HKEY_CURRENT_USER\SOFTWARE\ABBYY\FlexiCapture\12.0\Controller] "UseSpecialInvoiceLineItems"="false"

12. FlexiCapture for Invoices. Updated country settings for Netherlands
Reduced tax rate in Netherland changes from 6% to 9%. New tax rate is added to country settings.

13. Enable auto-learning for classification by script
Auto-learning capabilities became available for classification models that use a script.

14. Setup colors for document states
The color of document state for documents in the task can now be specified by user.

Known issue: the color setting is not translated into UI languages and remains in English.

Bug Fixes

Issue description
Custom project with invoices assembling has stopped working because Invoice Generic tried to merge all unknown pages, following the invoice, inside invoice document. The issues appeared in R2. Now the old behavior is recovered and Invoice Generic doesn't capture excess unknown pages.
If the multipage document was spitted at web verification, new documents could not be recognized.
Long tooltip with recognized text displayed over the image.
Optional footer inside Flexilayout behaved as obligatory footer, this caused wrong assembly of documents by Flexilayouts.
Line Items extrapolation did not work when using blob storage.
Fields conditions "Does not belong to set" and "Is not empty" in export settings could not be saved.
Mail Subject field was processed wrongly in utf-8 encoding using POP3, IMAP import from mail server.
Russian text about expiration date was displayed in License Manager run on other UI language after the first launch.

Issue description
An error occurred when getting a task assigned to a group with several users.
Some reports ended up with timeout.
Error occurred sometimes when moving a region of repeatable group.
Countries list in the FlexiCapture for Invoices help file is updated: countries with keywords only are listed separately.
Error 401 occurred during the connection if capital letter were used in tenant name.
The Document Definition was published after the upgrade of invoice project. Now invalid Document Definition will not publish.
Extra documents appeared when importing two files using AddNewDocument with excludeFromAutomaticAssembling=true.
User dictionaries did not work properly.
Repeatable groups training (including LineItems) did not work on multiple pages.
Fixed problems of vendor detection. Vendor is found correctly on some invoices where it was not found before o was found incorrectly.
An error occurred when merging documents with Unknown document definition on Web Verification Stations.
Some characters were incorrectly displayed on some UI languages in logs on the Administration and Monitoring Console.
Scroll of EULA was incorrect on Web Stations on small monitor.
An error occurred when user logged off without closing task on Web Verification Station.