ABBYY FlexiCapture 12 Release 3 Update 4 Release Notes

Table of Contents

Introduction

About this document
This document describes the changes made in ABBYY FlexiCapture 12 Release 3 Update 4.

About the product
ABBYY FlexiCapture 12 is the latest version of FlexiCapture, a stable and scalable Data Capture platform that can also be used for creating regional and vertical solutions.

About the patch
This update includes new features and bug fixes. It can be installed on its own, on top of ABBYY FlexiCapture 12 Release 1 or Release 2 (with any updates installed), or on top of ABBYY FlexiCapture 12 Release 3.

Technical Information

Release	Part #	Build #	Installer build #	OCRT build #	Release date
Release 3 Update 4	**1299/49**	**12.0.3.4164**	**318585**	**16.2.76.20**	**2021.04.09**
Release 3 Update 3 Patch 3	1299/48	12.0.3.4113	286686	16.1.1213.11	2021.02.24
Release 3 Update 3	1299/47	12.0.3.4093	253894	16.1.1213.11	2020.12.17
Release 3	1299/35	12.0.3.2525	70734	16.1.681.26	2019.10.04

New Functionality

1. ABBYY FlexiCapture for Invoices in P2P and O2C solutions

1.1. Extend country support of receipts.
The Neural Network for extraction fields from receipts was trained on new images from several countries, which allowed to extend country support and improve quality of extraction for supported countries.

- New countries supported in this update: Ireland, Guatemala, Panama, Honduras, Costa Rica, Uruguay, Argentina, Peru, Ecuador, Colombia, Brazil.
- Quality of extraction improved for: Spain, New Zealand, Australia, Turkey.

Quality tests

The table below shows the quality of extraction for header fields on receipts from the supported countries. For this benchmark, we used receipts, both photos and scans, from the main supported countries.

Country	Percentage with 3 correct fields	Date	Total	Currency	Country detection
Argentina**	69.60%	81.32%	82.05%	95.97%	96.7%
Australia	78.3%	85.71%	89.03%	96.45%	96.45%
Brazil**	61.30%	65.00%	92.50%	98.75%	98.75
Canada	86.1%	90.03%	90.77%	97.64%	96.59%
Colombia**	22.90%***	78.47%	39.58%	84.03%	100%
Costa Rica**	50.00%	75.47%	60.56%	96.27%	98.76%
Ecuador**	67.80%	76.90%	84.77%	97.54%	97.54%
Germany	83.3%	89.08%	95.23%	91.79%	93.89%
Guatemala**	70.00%	80.24%	78.98%	98.20%	98.8%
Finland	87.6%	92.34%	93.97%	99.39%	99.18%
France	86.5%	91.93%	92.93%	98.99%	97.97%
Honduras**	72.50%	84.71%	82.57%	97.86%	99.08%
Ireland**	82.10%	91.28%	89.74%	98.46%	98.46%
Italy	86.6%	92.38%	92.38%	99.6%	99.20%
New Zealand*	92.2%	94.95%	96.79%	98.62%	98.62%
Panama**	64.50%	90.44%	88.25%	77.32%	98.91%
Peru**	89.80%	92.67%	95.28%	97.88%	98.82%
Poland	82.1%	87.36%	91.95%	99.08%	99.54%
Spain	80.5%	91.18%	87.83%	97.00%	97.67%
Sweden	83.2%	93.88%	88.43%	98.97%	97.94%
Turkey	97.0%	97.3%	99.7%	100%	100%
UK	76.0%	86.09%	86.71%	98.77%	99.07%
Uruguay**	55.80%	86.04%	65.66%	98.11%	99.25%
USA	85.6%	91.95%	93.09%	96.80%	95.81%

Non-trained control countries					
Netherlands	88.6%	93.7%	93.30%	99.1%	98.61%
Singapore	74.3%	85.50%	84.73%	78.63%	93.13%

* New Zealand receipts can be processed starting from ABBYY FlexiCapture 12 Release 3 Update 3.

** Receipts of this country can be processed starting from ABBYY FlexiCapture 12 Release 3 Update 4.

***Low result of Total in Colombia is due to amounts without decimal parts. We will address this problem in next updates.

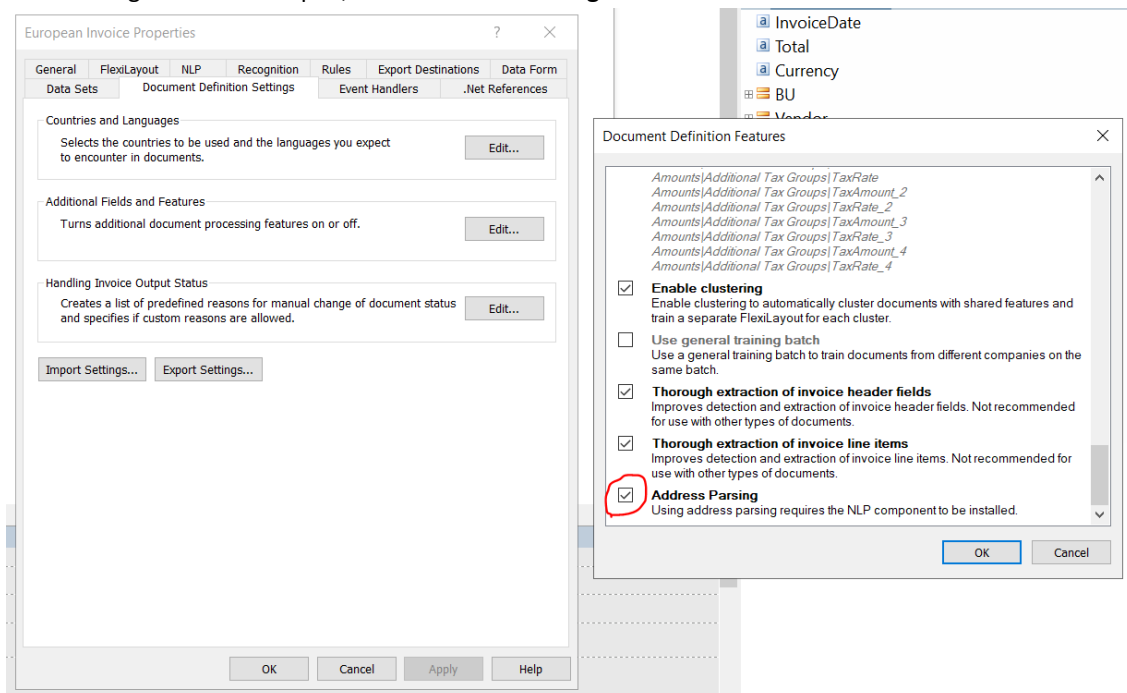# 1.2. Address parsing for invoices and purchase orders

Processing invoices and purchase orders requires extracting delivery addresses. To export them to the server system, the addresses must be split up into the following components as part of the format expected by the system: Postal code (ZIP), State, City, Street (Street, House, Apartment, P/O). From this release onwards, a new feature called **Address Parsing** has been implemented for FlexiCapture for Invoices. This feature allows you to parse the full address into different parts and can be used to process invoices and purchase orders.

Feature limitations:

1. Using the Address Parsing feature requires the NLP module to be installed.
2. Address Parsing is only available for English and German. If a language is not supported, the corresponding fields will not be filled in. However, the program detects the document language on its own instead of using the language specified in the settings. Therefore, if the program detects the language to be English or German, the address will still be parsed regardless of the language specified in the settings.
3. Enabling this feature may slow down document processing. The exact impact on processing speed will depend on the type of document and the complexity of addresses that need to be parsed.

By default, address parsing is disabled. To enable address parsing in new projects, do the following:

1. Navigate to the **Document Definition Settings** tab in the project properties.
2. In the **Additional Fields and Features** group, click **Edit…**
3. In the dialog box that will open, select **Address Parsing**.

Enabling this option will parse the Address fields for the Vendor and the BU. Processing these fields will export the region data for the following fields: Postal code (ZIP), State, City, Street (Street, House, Apartment, P/O).

To enable address parsing for projects created using earlier versions of the program, do the following:

1. Upgrade your project.
2. Enable **Address Parsing** in the **Document Definition Features** dialog.
3. Create a chain using the new **BU Address Parsing** rule for the **BU/Address** field and the **Business unit must be found in database** rule for the **BU** group.
4. Create a chain using the new **Vendor Address Parsing** rule for the **Vendor/Address** field and the **Vendor must be found in database** rule for the **Vendor** group.
5. Enable the **BU Address Parsing** and **Vendor Address Parsing** rules.

Note: If data sets are connected to the project, the detected address is first checked using the database.

- If an address is found in the database with high confidence, it is parsed from the database.
- If an address is not found in the database, or if it is found with low confidence, it is parsed from the image.

Changes to scripts

To support new script functionality, a new method called PutSpanToField has been added to the IFieldExtractor object. This method saves to the specified document field a text fragment that corresponds to the set text substring interval. From this release onwards, the PutSpanToField method will be used instead of SaveSpanToField. PutSpanToField differs from the old method in that it passes an IField object instead of a string as its last argument. For more information, please see IFieldExtractor.

Note:

1. Deprecated extraction methods will still be supported for existing projects.
2. For new projects, the use of deprecated methods will cause the batch to be sent to exceptions.
3. The new PutSpanToField methods can be used in all of the following: extraction scripts, rules, stages.

2. ABBYY FlexiCapture in enterprise environments

2.1. Security improvements

The following improvements have been made:

- Apache OpenOffice has been updated to version 4.1.9.
- Codemeter has been updated to version 7.10b.
- The following has been fixed: untrusted data could be used to redirect the user to a different website after login.

2.2. FCAdminTools: new SyncTrainingBatch command parameter

When working in several environments or with several projects at once, the **SyncTrainingBatches** FCAdminTools command is used in ABBYY FlexiCapture to transfer training results to other projects and/or environments. Sometimes the training batch created in the target project (in the production environment) needs to be deleted. In earlier versions, such batches would not have been deleted during training batch synchronization.

From FlexiCapture Release 3 Update 4 onwards, the **SyncTrainingBatches** will support a new parameter called **/DeleteBatchesMissedInSource**. When enabled, training batches that are not present in the source project will be deleted from the target project. For more information, please see SyncTrainingBatches.

2.3. Optimized database resource consumption

The following improvements have been made:

- OpenSession no longer puts an exclusive lock on the Session table.
- Calculation of statistics for training batches has been disabled.

3. ABBYY FlexiCapture web applications

3.1. Verification order mode hotkey for the Web Verification station

From this release onwards, the Web Verification Station now supports the verification order mode when pressing the Tab key. By default, the following order is used: top-down, left-to-right. This order can be changed in the Document Definition settings. For more information, please see the Selection order of fields section in Configuring how data is displayed in the document window.

3.2. Image color mode using URL-based API

A typical scenario for integrating the Verification Web Station component with third-party systems is as follows: an operator loads a batch of documents; the batch is processed and submitted for verification; an operator working in a third-party system is sent a link that opens the corresponding verification task on the Web Verification Station.

Such links are created using a URL-based API. From this release onwards, the color mode for images can be specified when opening a task on a Web Verification station. This is done using a new address bar argument called &loadInColor:
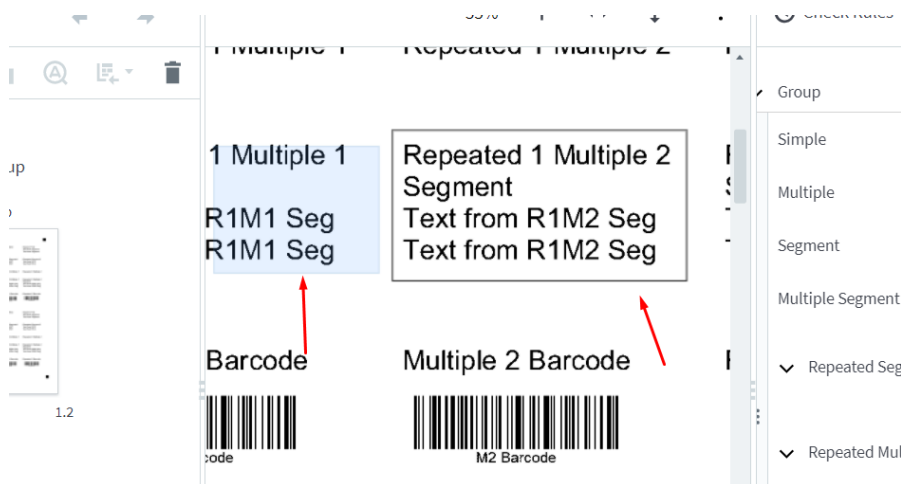
- &loadInColor=true will load the document image in full color on the verification station.
- &loadInColor=false will load the document image in black and white.

For more information, please see Integrating the verification stage.

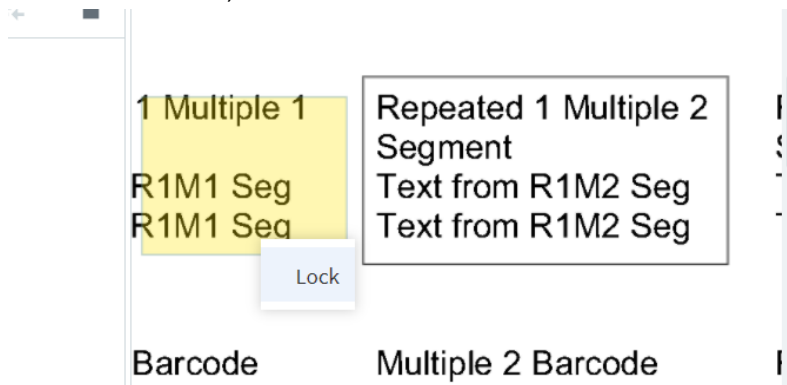## 3.3. Labelling segments for NLP processing

From this release onwards, Web Verification station operators will be able to mark up regions of other fields inside field regions. This allows you to label documents that will later be used to train segmentation and data extraction NLP models. For more information, please see Training your NLP models.

If Text segment has been selected in the field properties, the region for that field will not be filled in on the Web Verification station, allowing the operator to mark regions of other fields inside of it. Regions of fields that are not text segments are colored in blue and are locked by default, meaning that operators are not able to edit them.

To unlock a field, do the following:

1. Click the region of the appropriate field.
2. In the context menu, uncheck **Lock**.

Note: Once the **Check rules** command has been applied, all fields that are not text segments will become locked again.

3.4. New scanning plugin

From FlexiCapture 12 Release 3 Update 4 onwards, a new scanning plugin version will be available.

Important! Before installing the new scanning plugin version, uninstall the old version.

New plugin features:

1. The new plugin supports new neural network-based technologies such as barcode detection and using barcodes to split images into separate documents.
2. Allows you to send the source PDF file to the server in order to use its text layer for recognition. This option is set in the batch type settings and cannot be modified on a web station.
3. Administrators will be able to install the scanning plugin for all system users on a computer. This will allow administrators to set up shared working stations on which different operators are be able to work at different times. It is planned that future updates will allow operators to begin working on a Web Scanning station without requiring the scanning plugin to be installed. **Important**! The program does not support multiple operators working simultaneously on a single machine. Also, once the scanning plugin has been launched by a specific user, other users will not be able to scan documents. If this is the case, stop the scanning plugin process (ScanningPluginBR.exe) in Task Manager and then try to scan the document again.

Note. The size of the plugin has increased to the following values:

- 62 MB - plugin without PDF support
- 132 MB - plugin with PDF support

3.5. Proxy authentication support for the Web Scanning Station

The Application Server may be deployed within a protected environment, which may need the user to connect to it using a proxy that requires authentication. From this release onwards, a special proxy authentication mode will be implemented for the Web Scanning Station.

Using this feature requires the following:

1. The new scanning plugin should be installed.
2. An **IsScanPluginProxyRequired=true** parameter should be added to the "**AppSettings**" section of the Web.config file (the default location of this file is c:\inetpub\wwwroot\FlexiCapture12\Monitoring).

Connecting to the server for the first time will bring up a dialog where the user can input his or her credentials. If the authentication is successful, the user details are saved in the registry, meaning that the user will not have to provide the credentials a second time.

4. General improvements

4.1. Changes to the default settings of the Auto mode for processing PDF files
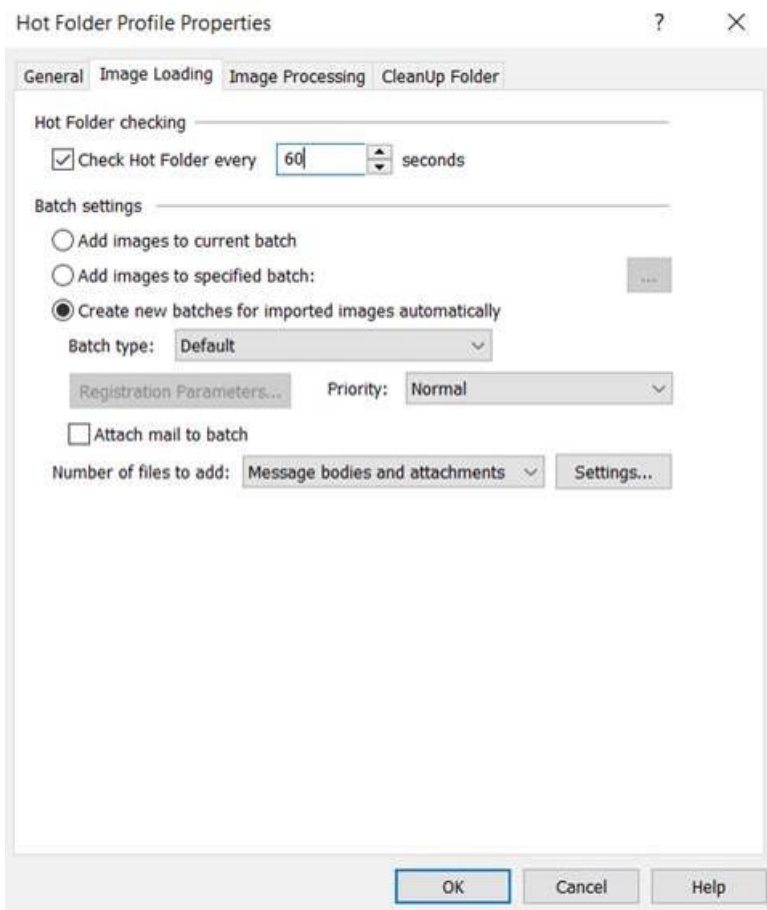
A PDF document may contain both a text layer and raster images (e.g. scanned pictures or embedded text). In FlexiCapture Release 3 Update 2, a new recognition mode was added specifically for these PDF document types. In this mode, the program takes the text directly from the text layer (without performing any OCR) and recognizes the raster objects. Should the text layer and a raster object overlap, the text layer is preferred. Previously, this mode had to be enabled using a registry key. From this release onwards, this mode will be enabled by default as long as **Auto mode** is selected in the PDF processing settings.

You can revert to the legacy recognition algorithm by modifying the following registry keys:

- For FlexiCapture: [Computer\HKEY_CURRENT_USER\Software\ABBYY\FlexiCapture\12.0\DAForms] EnablePullXTextAndRecognizeRest=**false**
- For FlexiLayout Studio: [Computer\HKEY_CURRENT_USER\Software\ABBYY\FlexiLayoutStudio\12.0\DAForms] EnablePullXTextAndRecognizeRest=**false**

4.2. Improved automatic import of e-mails from mail boxes

Before this release, using automatic mail box import imported one e-mail message at a time and did not import any more messages prior to the expiration of the **Check Hot Folder every N seconds** time period. This patch changes this behavior as follows. After the first e-mail message has been imported, the program will immediately check if there are any more messages in the mail box. If more messages are detected, they will be imported, one by one, until the mail box is empty. Once the mail box is empty, the program will begin checking the mail box at regular intervals specified in the **Check Hot Folder every N seconds** box.

Hot Folder Profile Properties

General | **Image Loading** | **Image Processing** | **CleanUp Folder**

Hot Folder checking

☑ Check Hot Folder every [60] ⇕ seconds

Batch settings

○ Add images to current batch

○ Add images to specified batch: [...]

◉ Create new batches for imported images automatically

Batch type: [Default ⌄]

[Registration Parameters...] Priority: [Normal ⌄]

☐ Attach mail to batch

Number of files to add: [Message bodies and attachments ⌄] [Settings...]

[OK] [Cancel] [Help]

4.3. Offline scanning station functionality

Scanning operators may be required to travel out of the office, for example, to help a client fill out a finance application. This means that the operator may not always have access to the internet, and as a result, the FlexiCapture servers. However, the operator may still be required to scan and carry out initial processing for documents (e.g. process images, fill out registration parameters). To help deal with such cases, an offline mode for the scanning station has been introduced in this update.

The following features are available in the offline mode:

- Scanner management: start and stop the scanning process.
- Set up scanning parameters, batch parameters, page size, etc.
- Prepare images for further processing: process images in the Image Editor, specify registration parameters.

Using the Scanning Station in offline mode has the following limitations:

- Documents cannot be exported.
- Batch types cannot be managed.
- You can only use batch types that have been created in advance using a connection to the FlexiCapture server.

Access to all of the program's features requires an internet connection to the Processing Server.

4.4. New registration parameter

When importing images from mail boxes (using either the POP3 server protocol, IMAP, or MAPI), FlexiCapture automatically records addressees from the To field to the **fc_Predefined:EmailAddressee** registration parameter of the created batch. However, the CC field can also contain an address or address group that may affect the way the document is processed. From the current release onwards, the **fc_Predefined:EmailCopyAddressee** registration parameter has been implemented in FlexiCapture to deal with such cases. This parameter is used to record all addresses specified in the CC field.

Note: The maximum length of text for this parameter is 255 characters. If the total number of characters in the CC address exceeds 255, all remaining characters will be omitted.

Known Issues

Installing this update via the command line will cause the WIBU_USER_CMD=No key to not work.

Bug Fixes

Bug ID	Issue description
Zendesk#321998	The Continue LineItems function (not designed for table elements) was available in the UI for table elements.
Zendesk#352192	An incorrect default language was selected when creating a document definition in Chinese.
Zendesk#184385	A field region was detected incorrectly after training using a sufficient number of samples.
Zendesk#197096	The Environment Variable dialog could not be sorted.
Zendesk#304265, Zendesk#328073	The Web Scanning Station and Web Capture Station didn't send the PDF text layer to the server, meaning that it was not used during recognition.
Zendesk#331867	The Entry Condition dialog displayed incorrect values when saving changes and re-opening.
Zendesk#342526, Zendesk#358888, Zendesk#388524	Vendors that were not relevant for the selected BU were displayed in Vendor Lookup.
Zendesk#342349	Sometimes text was displayed as a sequence of dots when importing PDF files with a text layer.
Zendesk#361179, Zendesk#357490	On the Web Verification station, deleting all text from a field when correcting errors caused the field to change its size and become unresponsive.
Zendesk#347293, Zendesk#336943	The "@" symbol could not be entered on the Web Verification Station using the German keyboard layout.
Zendesk#306920	Different names for the same feature in the English and Russian UI's.
Zendesk#347703 Zendesk#378948	Changing the Ratio property via the shell scripts did not rescale the window.
Zendesk#343992, Zendesk#371822, Zendesk#391389	The PerMachine Web Scanning Plugin installation option did not work as intended.

Bug ID	Issue description
Zendesk#367691, Zendesk#343042	A 302 Redirect occurred when opening iFrame in Chrome.
Zendesk#355128	Different results for full-text and block recognition.
Zendesk#359450	An "Src\ProductItemsFinder.cpp, 856" error occurred when trying to recognize Hebrew invoices.
Zendesk#367320	An "Src\FormAnalysis.cpp, 1952" error occurred when checking a document definition, caused by incorrect mapping between DataSet columns and the database.
Zendesk#327854	Using the Duplicate command in the invoice project settings caused some countries to be missing from the Business Unit country list.
Zendesk#318306	An "Exit code: 4" error occurred in FlexiCapture 12 when opening a Hot Folder file using the embedded Office application.
Zendesk#286535	A "Scanner not found (Can't locate device; check cable and power.(-4536)" error occurred when working with some scanners on the Web Scanning station.
Zendesk#362551	Filtering by classes on the Project Setup Station sometimes returned incorrect results.
Zendesk#356567	Several document sets containing the same nested document caused the incorrect document to be chosen even if the available document sets were restricted by batch type.
Zendesk#347729	A "DatabaseCache.Query.h, 238" error occurred when trying to look up a record in DataSet on the Web Verification Station.
Zendesk#205501	The Inflate property was incorrectly passed to the child block.
Zendesk#370773	Black square anchors were not detected correctly on some documents.
Zendesk#348949, Zendesk#377423, Zendesk#270261	Some rules were checked only once on the Web Verification station.
Zendesk#349491	Training sometimes did not improve extraction results for projects.
Zendesk#371085	Some PDF files could not be opened.
Zendesk#373421	Repeating fields were not shown in the Web Station if there were multiple repeating fields already on the form.
Zendesk#313193	A "Contains invalid path" error occurred when updating a Document Definition.
Zendesk#366379	Upper-case text was not recognized in some PDF files with a text layer.
Zendesk#349586	The registration parameter list did not work as intended if there were two or more values starting with the same string of characters.
Zendesk#382004, Zendesk#383651	WDSL validation errors caused by Web API responses.
Zendesk#330091, Zendesk#386447, Zendesk#390457, Zendesk#390105	Sometimes PDF image layouts got corrupted when imported to FlexiCapture.
Zendesk#351425	Text near a stamp was not recognized on some documents.
Zendesk#378819	No error message was displayed when upgrading a project with one or several corrupted batches, making it appear that the upgrade has been carried out successfully.

Bug ID	Issue description
Zendesk#382840	A "SparseFloatMatrix.cpp, 129" error occurred when training NLP segments on some documents.
Zendesk#384956	A "fetchDocumentDataError" error occurred on the Web Verification station when an invoice was sent for processing.
Zendesk#352835	Sometimes importing of multi-page documents from e-mail was interrupted.
Zendesk#377395	A "Can't Reach This Page http://localhost:32023" error occurred when setting up MS Graph import.
Zendesk#381967	XLS files containing a large number of columns were split into several pages.
Zendesk#350053	Connecting to existing SQL databases using Windows authentication was not possible.
Zendesk#385554	A barcode.dll module crashed during image processing.
Zendesk#338439	Using a script to check all rules did not work as intended.
Zendesk#214344	The RecognizeText method did not always work in FlexiLayout Studio.
Zendesk#390831	It was not possible to scroll from one page of the document to the next after selecting "Fit-To-Width" on the Web Verification station.
Zendesk#378291	Gregorian calendar dates were not converted correctly into Buddhist calendar dates.
Zendesk#384211	A "the field value is not a date" error occurred for the Spanish date "2 de Jun del 2020" (the Spanish language and the "May include month in words" were both enabled).
Zendesk#377584	It was not possible to delete characters represented by surrogate pairs using the backspace key on the Web Verification station.
Zendesk#388368, Zendesk#397645, Zendesk#397663, Zendesk#396718, Zendesk#397683, Zendesk#393828, Zendesk#397753, Zendesk#386739	An "Object reference not set to an instance of an object" error occurred when importing via MS Graph.
Zendesk#388209, Zendesk#392221	A "Code: generalException Message: An error occurred sending the request." error occurred when trying to set up MS Graph with TLS1.0/1.1 disabled.
Zendesk#388772	The document form could not be refreshed automatically (except by scrolling manually) after selecting a document with another definition during task verification on the Web station.
Zendesk#371546	Updating the document definition version for a training batch document caused documents to switch to the Unused state.
Zendesk#383115	Recognition results after training were not correct for some projects.
Zendesk#390145	Several radio buttons could be selected in the FlexiLayout Studio UI (Page Import Setting dialog).
Zendesk#359437	Web Stations were inaccessible in WebSeal environments due to incorrect URL parsing.
Zendesk#385108	Data was not being saved on the Data Verification Station when the station window was closed.
Zendesk#346690	A "Not enough memory!" error occurred when trying to download large projects from a server.

Bug ID	Issue description
Zendesk#368830, Zendesk#365179, Zendesk#373160	High system loads caused the available pages counter in the Administrator's and Monitoring Console to be updated with delays.
Zendesk#381384	DataSet.GetRecords(query) returned incorrect results on some invoice projects.
Zendesk#392725	Source batch comments were not copied to the target batch when synchronizing training batches using FCAdminTools.
Zendesk#382425	Japanese text was not recognized in Generic Invoice Layout in some Japanese projects.
Zendesk#305074, Zendesk#190661	Recognition quality for CJK text was improved.
Zendesk#377247	The Group into Section option was not localized.
Zendesk#389038	PDF417 barcodes were not being detected on some images.
Zendesk#340522	Some PDF files were being imported into FlexiCapture as blank pages.
Zendesk#321998	The Continue Line Items command was available for an incorrect block type (Table), the use of which caused a "FineObjects\lnc\Object.h, 344" error.
Zendesk#379023	The Chinese help files contained incorrect product names.
Zendesk#400594, Zendesk#400675, Zendesk#402466	A "DataTables warning ..." message appeared when opening the Processing Monitoring tab on the Administrator's and Monitoring Console.
Zendesk#342818	Field regions had to be drawn/selected twice in Field Extraction Training.
Zendesk#400620	An "Unhandled Internal Error" error occurred in FlexiLayout Studio when working with a FirstFound element.
Zendesk#391025	Changing the size of the document thumbnail panel during a task sometimes caused the screen to become blank.
Zendesk#363430, Zendesk#372566	Purchase order numbers were not detected correctly in some US Invoice projects.
Zendesk#326783	Certain header keywords could not be detected on some images.
Zendesk#336069	The hyphen character was not recognized correctly on some images.
Zendesk#373025	Parts of text could not be captured on some documents.
Zendesk#387885	The JPY currency was not set as the default option in the Japanese Invoice Project.
Zendesk#403055	The Total field was not detected correctly on some invoices.
Zendesk#386269	It was not possible to create a multi-region field on the Web Verification Station.