ABBYY FlexiCapture Engine 12 Release 2 Update 3 Release Notes

Table of Contents

1 About the Document

This document contains a description of Release 2 Update 3 (R2 U3) of the ABBYY FlexiCapture Engine 12.

The *Release Notes* is a document intended to be the source of technical information about the Release.

2 About the Release

2.1 Purpose of the Release

Purpose of this release is bugfix. Please find an additional information in the "Fixed Bugs" section.

2.2 Product Information

ABBYY FlexiCapture Engine 12 is a new release of the ABBYY Data Capture SDK delivering new features and technology improvements.

ABBYY FlexiCapture Engine is a development tool for ISVs, OEM vendors, and service providers that wish to develop data capture solutions for processing fixed-form documents and semi-structured documents with repetitive layouts. It can also be used for creating customized data capture applications for various scanning devices and terminals.

2.3 Release Description

This is a release for worldwide sales.

2.4 Part Number, Build Number

Part # 1355/19

Build # 12.1.25.9

2.5 Technology and FlexiCapture Build Numbers

FlexiCapture	12.0.3.4066
OCR Technology build #	16.1.1014.16
APDFL version	15.0.4PlusP5e

ABBYY FlexiCapture Engine 12 works with projects created in ABBYY FlexiCapture 12, Build 12.0.3.4066, or earlier. Projects saved in the current version of ABBYY FlexiCapture Engine 12 can be used in ABBYY FlexiCapture 12, Build 12.0.3.4066, or later. You can also pass projects to and from ABBYY FlexiCapture 12 Developer's Package shipped with the current ABBYY FlexiCapture Engine 12 distribution.

2.6 Software and Hardware Requirements

2.6.1 ABBYY FlexiCapture Engine 12 Requirements

- PC with x86-compatible processor (1 GHz or higher)
- Operating system (compatible with last available service packs):

  - Windows Server 2016, 2012, 2008 R2
  - Windows 10, 8.1, 8, 7
- ABBYY FlexiCapture Engine has been tested on the following cloud computing platforms:
  - Microsoft Azure App Services, Cloud Services, Service Fabric, Virtual Machines
  - Amazon EC2
- ABBYY FlexiCapture Engine has been tested in the following virtual environments (tested on last available service packs):
  - Microsoft Hyper-V Server 2016, 2012, 2008
  - Oracle VM VirtualBox 5.2
  - Parallels Desktop for Mac 13.0.1
  - VMware ESXi 6.5
  - VMware Workstation Player 12.5
  - VMware Workstation Pro 14.0.0

    ABBYY FlexiCapture Engine can also be run in a Docker container on supported platforms.
- Memory:
  - For processing one-page documents — minimum 400 MB RAM, recommended 1 GB RAM
  - For processing multi-page documents — minimum 1 GB RAM, recommended 1.5 GB RAM
- Hard disk space: 1400 MB for library installation and 100 MB for program operation plus additional 15 MB for each processed page of a multi-page document
- 100% TWAIN-compatible scanner, digital camera, or fax modem — for scanning only
- Video card and monitor (min. resolution 1024*768) — for displaying GUI elements
- Keyboard, mouse or other input device
- The following registry branches should be accessible from the workstation:
  - "HKEY_CURRENT_USER\Software\ABBYY\SDK\12\FlexiCapture Engine" — full control
  - "HKEY_CURRENT_USER\Software\ABBYY\SDK\12" — full control for installation only
  - "HKEY_LOCAL_MACHINE\Software\ABBYY\SDK\12" — full control for installation only
- The following folders should be accessible from the workstation:
  - Folder with ABBYY FlexiCapture Engine binary files — access for reading
  - %TEMP% folder — full control access
  - Folder %ProgramData%\ABBYY\SDK\12\FlexiCapture Engine — full control access
  - Folder %ProgramData%\ABBYY\SDK\12\Licenses — full control access
- The following components should be installed:
  - Microsoft Internet Explorer 8.0 or higher

2.6.2 ABBYY SDK 12 License Server Requirements

- PC with x86-compatible processor (1 GHz or higher)
- Operating system (compatible with last available service packs):
  - Windows Server 2016, 2012, 2008 R2
  - Windows 10, 8.1, 8, 7
- ABBYY FlexiCapture Engine has been tested in the following virtual environments (tested on last available service packs):
  - Microsoft Hyper-V Server 2016, 2012, 2008
  - Oracle VM VirtualBox 5.2
  - Parallels Desktop for Mac 13.0.1
  - VMware ESXi 6.5
  - VMware Workstation Player 12.5
  - VMware Workstation Pro 14.0.0

    ABBYY FlexiCapture Engine can also be run in a Docker container on supported

- 25 MB of free hard-disk space
- Folder %ProgramData%\ABBYY\SDK\12\Licenses — full control access
- For activation/deactivation of the license:
  - Allowed connections to *.abbyy.com
  - Port: 80
  - Protocol: HTTP

  If proxy-server, VPN or dial-up connection is used, you must customize the following settings in Internet Explorer: *Tools* > *Internet Options* > *Connections*

2.7 Upgrading from Previous Versions and Releases

2.7.1 Binary Incompatibility

It is necessary to recompile the host application regardless of the version of the Engine previously used.

2.7.2 API Changes

2.7.2.1 Release 1

Please refer to the "ABBYY FlexiCapture Engine 12 and 11 Compatibility" section in the product documentation.

The main changes are related to:

- project structure
- document definition structure
- image-related objects
- scanning-related objects
- license-related objects

2.7.2.2 Release 2

2.7.2.2.1 New algorithm for FCE loading

The loading of the Engine object is now a unified algorithm with the possibility to customize the loading parameters. For details, see *Guided Tour* > *Tutorial* > *Processing with a Customizable Processor Object* > *Step 1: Load FlexiCapture Engine* in the User's Guide.

2.7.2.2.2 Automatic text- and image-based classification is deprecated

FlexiCapture Engine is designed for data extraction from fixed forms and semi-structured documents. Normally, these types of documents can be described by using rules. There is no need to set up an automatic classifier to classify such documents.

The corresponding functionality will be removed in future versions.

2.7.2.2.3 Working with projects is deprecated

As FlexiCapture Engine is intended for processing fixed forms and semi-structured documents, there is no need to use complex FlexiCapture project interfaces. These scenarios can be implemented using the light-weight FlexiCapture Processor.

The corresponding functionality will be removed in future versions.

2.7.2.2.4 ActiveX visual components are marked as deprecated

Statistics show that the ActiveX components are almost never used in the product. For this reason, it was decided to mark this functionality as deprecated.

The ActiveX components will be removed from the product in the next version.

2.7.3 Changes in Behavior

2.7.3.1 Release 1

2.7.3.1.1 New default installation folders

Compared to the previous version, this release has different installation folders.

Folder	*Description*
"%ProgramFiles%\ABBYY SDK\12\FlexiCapture Engine"	This is the **default** folder for storing binaries, the Help file, the guides, and the USB dongle driver redistribution. The Read Me file is also placed here. The user can change this destination folder when installing the product.
"%ProgramData%\ABBYY\SDK\12\Licenses"	This is a permanent place for storing the protection subsystem files.
"%ProgramData%\ABBYY\SDK\12\FlexiCapture Engine"	This is a permanent place for storing the auxiliary Engine files, such as the *include* files and samples.

3 List of Changes

3.1 New Features and Improvements

3.1.1 Release 1

3.1.1.1 Technologies

Compatibility with FlexiCapture 12 – the product should be based on the same technologies (v. 15 technologies).

The new technologies bring several new features to the product. These features will be available in ABBYY FlexiCapture Engine 12 Release 1 only via Document Definitions or other FlexiCapture entities (e.g. classifiers) created in ABBYY FlexiCapture 12. They will be available through the FlexiCapture Engine API in later releases.

3.1.1.1.1 Export

A new export setting has been added to create files for archiving in back-end systems with specific requirements for file formats, versions, and properties.

In the case of export to PDF, FlexiCapture offers a combination of settings that strikes a balance between file size and image quality. However, users can change the default settings to get files of minimal size or images of maximum quality.

- Users can now select a version of PDF/A. The list of available versions depends on the selected standard.
- The quality setting allows selecting between file size and image quality. Use "low" if you need the smallest possible files, use "high" if you need the best quality, or use "normal" for a balance between the two.

- MRC can now be used for non-searchable PDFs. Use this setting to get compressed PDFs that are small in size but have good visual quality.
- The “Fast Web View” option allows showing the first page of a multi-page document before the entire document has been loaded. Use this setting for documents that will be put up on web pages. This setting is ignored for one-page documents.
- Use the “Enable tagged PDF” setting for searchable PDFs if you need to add tags.
- A compression codec can now be selected for color and grayscale images exported in TIFF. Select “Custom quality” and then select a desired compression algorithm.

UTF encoding without BOM is now supported for all text formats (CSV, TXT, and XML).

3.1.1.1.2 Miscellaneous

3.1.1.1.2.1 Working with repeating fields and groups

In the previous versions of FlexiCapture, a repeating group could only be created in ABBYY FlexiLayout Studio. Now the Document Definition Editor allows creating repeating fields and groups, and can be used for training or for manual filling.

3.1.1.1.2.2 Using additional FlexiLayouts for a Document Definition

One Document Definition can now work with multiple FlexiLayouts. Among these, one is a main FlexiLayout and the other are additional FlexiLayouts. The main FlexiLayout is responsible for document identification and defining the document borders. The additional FlexiLayouts will only be used if the main FlexiLayout is matched. Additional FlexiLayouts can add more fields or re-define already matched fields. The following rule is used to combine fields from different FlexiLayouts: each subsequent FlexiLayout will either add a field with a unique name, or overwrite a field’s region if a match was found in one of the previous FlexiLayouts. You can organize communication among FlexiLayouts within one Document Definition using named parameters.

3.1.1.2 Protection and licensing

Customers working with several Developer serial numbers, use their own Developer serial number for FCE initialization instead of the Master Developer serial number (Project ID). This causes errors when the program is used by end-users.

The behavior of FCE has changed. We have introduced a new entity, “Customer Project ID.” This entity has the following important characteristics:

- It is not a serial number — just an ID!
- The Engine must be initialized using a Customer Project ID.
- All serial numbers are connected to it.
- It’s format is different from the serial number format (e.g. F3O1OV0CkT1PTXIa3dyp).

If FCE is initialized with any text string other than a Customer Project ID, the compiled program will generate an exception saying “Failed to run license <…>. Incorrect Customer Project Id.” FCE works as follows:

Serial number, used in the source code for FCE initialization / License, activated on the work station	Customer Project ID	Any Developer SN	Any Runtime SN
Developer license	+	-	-
Runtime license	+	-	-

3.1.2 Release 2

3.1.2.1 Improved speed and accuracy, new languages, and barcode types

The new OCRT 16 technologies give developers the opportunity to use the latest available OCR technologies in their solutions.

In addition to speed and accuracy improvements, the new technologies allow using new languages and barcode types:

- MaxiCode barcode type.
- Japanese Modern language support (consists of Japanese characters, full set of English characters, and 4 Greek characters: α, β, θ, π). Recommended for modern Japanese documents.
- Burmese, Farsi, and Georgian recognition language support.

3.1.2.2 New filters to enhance images manually

New image enhancement methods are now available in the product. They are available through the properties of the ImageLoadingParams object.

The new methods allow users to:

- Reduce noise in high-ISO images
- Remove color marks, such as stamps or signatures made with colored ink, from the image
- Automatically crop images:
  - detect image boundaries
  - crop area which contains data to be extracted
  - correct skew and distortions
- Whiten image backgrounds

It is also possible to treat a preprocessed image as a photo, applying filters specific to photo images.

All these new filters allow implementing solutions which automatically improve images captured with mobile devices and address complex backgrounds of documents like IDs, insurance policies, etc.

3.1.2.3 Digital PDF processing

Sometimes users need to extract data from PDF files which contain a text layer. Using information from this layer may increase the accuracy of PDF recognition and subsequent data extraction. It can also speed up document processing.

Earlier, when loading a PDF file, FCE used only image information from the file. It is now possible to select a mode for reusing the content of the source PDF file. The following three modes are available:

- **Content only**. In this mode, only the text layer of the source PDF file will be used for data extraction. The image layer will not be used. This mode is of no use if the source file contains only raster information (e.g. if the source file is an image-only PDF). In this case, FCE may use OCR to improve data extraction accuracy.
- **Do not reuse**. In this mode, the text layer of the source PDF file will not be used for data extraction. The image layer will be recognized by FCE using OCR.
- **Auto**. In this mode, FCE will use both the text and image layers of the source PDF file. This mode is selected by default.

These modes can be selected using the SourceContentReuseMode property of the ImageLoadingParams object.

3.1.2.4 New export options

- The ability to select PDF/A version during export
  - Earlier, it was possible to create only PDF/A-1a documents in FCE. Some organizations may need to use different subtypes of PDF/A format to comply with legal requirements. It is now possible to choose any subtype of PDF/A for export: PDF/A-1a, PDF/A-1b, PDF/A-2a, PDF/A-2b, PDF/A-2u, PDF/A-3a, PDF/A-3b, or PDF/A-3u.
- "Fast Web View" option for export to PDF
  - With the "Fast Web View" option selected, users will be able to start reading a PDF without waiting for the download to complete. This is especially useful for large documents which take a long time to download.
- The ability to export tagged PDFs
  - A tagged PDF is a structured PDF that allows page content to be extracted and used for various purposes, such as reflow of text and graphics, conversion to HTML and XML, and making content accessible to the visually impaired.
- The ability not to write a BOM character when exporting to UTF.
  - Earlier, FCE always added a Byte Order Mark (BOM) when exporting to TXT. UTF-8 permits the BOM but does not require or recommend its use. Java's UTF-8 encoding does not recognize this character as a BOM. As a result, Java treats the BOM character as garbage.

3.1.2.5 FRE 12 API included

FCE delivers fast-response tools for data capture, while ABBYY FineReader Engine is ideal for archiving applications to produce highly compressed images, searchable PDF files, and documents that retain all of the original layout and formatting. For this reason, customers may want to use features of both ABBYY FineReader Engine and ABBYY FlexiCapture Engine in their applications.

For this purpose, FCE provides an easy way of integrating it with FRE. Moreover, the ImageDocument object of FCE may be transferred to FRE, and vice versa.

To use the FRE 12 API inside FCE 12, a special license may be required.

3.1.2.6 New engine deployment options

Cloud environments are infrastructures and services that provide Internet access to shared resources (e.g. servers, data stores or applications). Cloud environments are often based on virtual machines, which are deployed for each application in the cloud.

Cloud environments are becoming increasingly popular, as more and more customers are deploying their solutions in such services as Amazon EC and MS Azure. FCE 12 now supports a new licensing type, which allows running the product in virtual and cloud environments.

3.1.2.7 64-bit version

The terms “32-bit” and “64-bit” refer to the way a computer’s CPU handles information. A 64-bit operating system handles large amounts of random access memory (RAM) more effectively than a 32-bit system.

FCE 12 now natively supports 64-bit operating systems.

3.1.3 Release 2 Update 2

3.1.3.1 A single method to correct image orientation

FlexiCapture Engine has several methods for images preprocessing before recognition: *DetectOrientationByText* and *RotateImageByRotationType*. The first one detects orientation of an image by text lines inside it. The second one rotates an image by a specified rotation: 90 degrees clockwise, 90 degrees counterclockwise or upside down.

Developers, who needs to rotate an image to a normal orientation, should use these two methods sequentially. It was not comfortable for the customers. Because they needed to get a rotation of the image as result of the first method. Then to invert rotation to get a rotation which must be applied to the image to return it to the normal orientation. And finally, to call the *RotateImageByRotationType* method.

Now it is possible to call a single method *CorrectOrientationByText*, which will do the described sequence of actions automatically.

3.2 Fixed Bugs

This section contains a list of bugs reported by customers that have been fixed in the product.

The following error classification will help you to evaluate the severity of each issue and make an informed decision about how important the updates are for your system.

Critical	A bug that causes crashes or freezes. Critical bugs include access violations, internal program errors, stack overflow, out-of-memory or other exceptions that can lead to program failure.
Major	A bug that does not cause program failure but affects major functionality of a feature or impairs the system’s performance. Major bugs include disparity between the actual operation of a feature and the internal specifications, memory leaks or data corruption.
Minor	A bug that leads to feature malfunctioning or affects minor software features. Minor bugs include recognition errors, missing or lost objects, wrong color detection, incorrect document analysis, license counter errors, etc.
Trivial	A cosmetic issue that does not affect the functionality of the product but may cause inconvenience. Trivial bugs include Help file errors, log errors, incomplete information in error messages, etc.

Severity	**Description**	**HD or ZD #**
Critical	AV error when using ApplyFrequencyTransform Method of the ImageProcessingTools Object on Java. **Solution**: the filter is marked as deprecated, imageProcessingTools should be used instead.	305059

	Workaround: change NumberOfLevels from 10 to 9 in the transformation parameters.	
Minor	When working with Online license, check for the certificate revocation is failed. It leads to inability of the licensing to work	259593
Minor	"License parameters do not match required parameter" error, when FCE switches from umbers, generated for FCE, to serial number, generated for FCE with FRE API. As a **workaround**, customer may deactivate license without FRE API in case it is necessary to use license with FRE API on the same machine with the same customer project ID.	230725
Minor	IFieldValue::AsDateTime returns wrong values for dates before January 1st, 1970 in Java	220101
Trivial	IImageProcessingTools::CalcSuitableForOcr() method is described in the product documentation while is missing in the product.	246202

