

PART NUMBER, BUILD NUMBER

Part#	1041/20
Build#	11.1.5.104

NEW FEATURES AND IMPROVEMENTS

NEW SAMPLES FOR JAVA

New samples are available for Java:

- BatchProcessing

This sample shows how to use Batch Processor for processing a large amount of one-page documents.

- EventsHandling

The sample illustrates the use of the callback interfaces using the FRDocument callback interface (IFRDocumentEvents) as an example. The sample shows the progress of recognition and export during image processing. You can use the callback interfaces to control image processing.

- EnginesPool

This sample on the one hand provides a complete reusable solution for a pool of Engines in a multithreaded application, and on the other hand demonstrates the gain in speed when using multiprocessing.

The samples have the same scenarios as the samples with the corresponding title provided in C#.

IMPROVED JAVA WRAPPER

In previous release Java wrapper had some limitations. Now the complete version of JAVA wrapper is implemented, which means that entire SDK functionality is available through it. Now FRE can be easily integrated into Java environment.

New features are supported:

1. Inheritance of interfaces
If interfaces are inherited in IDL, the corresponding wrapper interfaces are also inherited (For example, ITextBlock and IBlock)
2. Methods that load data from memory(e.g. IFRDocument.AddImageFileFromMemory)
3. Interfaces which are implemented on the client side (callbacks and events) :
 - a. The IExternalDictionary interface
 - b. The IAsyncProcessingCallback, IImageSource, IFileAdapter interfaces
 - c. The IDocumentAnalyzerEvents interface

- d. The IExporterEvents interface
 - e. The IFRDocumentEvents interface
 - f. The IFRPageEvents interface
 - g. The IFRPagesEvents interface
 - h. The IImageDocumentEvents interface
 - i. The IImagePasswordCallback interface
 - j. The IScanCallback interface
4. Methods with out-parameters (For example, IPlainText::GetCharacterData method)

ENGINE::INJECTTEXTLAYER METHOD WAS ADDED

This is a new method which creates a copy of the input PDF file and adds the text layer which corresponds to the recognized text of the document.

SUPPORT FOR INTELLIGENTMAIL BARCODES

Starting with this release, ABBYY FineReader Engine can recognize ABBYY FineReader Engine recognizes IntelligentMail barcodes. Intelligent Mail is a height-modulated 65-bar barcode which is used on mail in the United States. It is also known as USPS 4-CB.

BT_IntelligentMail was added to BarcodeTypeEnum enumeration constants.

PREDEFINED LANGUAGES OCRA AND OCRB WERE ADDED

The OCR A and OCR B modules now license not only corresponding text types but also the predefined languages which are provided for these text types. Apart from regular characters the alphabet for these languages also contains special symbols from OCR_A and OCR_B standards.

NEW SPECIAL PREDEFINED LANGUAGES FOR PHONE NUMBERS AND EIN

English_US_Address_PhoneNumber and English_US_WellKnownCode_EIN special languages were added.

English_US_Address_PhoneNumber contains phone numbers (English, United States).

English_US_WellKnownCode_EIN contains Employer Identification Numbers (English, United States).

DETECTION OF VERTICAL TEXT FOR ALL LANGUAGES

IPageAnalysisParams::DetectVerticalEuropeanText property is a new property which allows you to detect vertical text in languages other than CJK (helpdesk request 369221).

NEW ICLASSIFICATIONTRAINER::ADDPAGE METHOD

From now on FRE 11 supports IClassificationTrainer::AddPage method.

This is a new method which allows you to add an already recognized page to the classification database.

THE POSSIBILITY TO ADD CUSTOM FEATURES FOR CLASSIFICATION

Now it is possible to add custom features that can enhance the accuracy of classification.

The following methods were implemented:

1. `IClassificationTrainer::AddFeaturesForPage`

This is a new method which supports training the classification database with the help of user-defined additional features.

2. `IFRPage::ClassifyEx`

This is a new method which supports classification with the help of user-defined additional features.

NEW PREDEFINED PROFILE ENGINEERINGDRAWINGSPROCESSING

This profile is specially developed for engineering drawings and maps recognition.

`EngineeringDrawingsProcessing` would be useful for text extraction disregarding its orientation and conversion of engineering drawings from PDF to searchable PDF.

EXPORT WITH ORIGINAL LAYOUT TO XLSX

New property `IXLExportParams::LayoutRetentionMode` that allows to export files to XLSX with original layout was added.

It is possible to choose the mode of retaining tables' formatting during export to XLSX. `XLSXLayoutRetentionModeEnum` gives the ability to choose the mode that is better corresponding to the assigned task.

NEW PROPERTIES FOR IBARCODEPARAMS OBJECT

This release includes a new property which allows you to detect more barcodes but slows down the processing. `IBarcodeParams::EnableAdvancedExtractionMode` property enables the feature.

New `IBarcodeParams::MinRatioToTextHeight` property defines the minimal acceptable height of the barcode in relation to the average letters height. This setting can help to detect low barcodes.

SUPPORT FOR EMBEDDED FILES IN PDF

`IPDFExportFeatures::WriteSourceAttachments` property specifies if attachments from the original PDF file should be written to the output file. This property is ignored if the original file was not PDF.

New values `PCM_Pdfa_3a` and `PCM_Pdfa_3u` were added to enumeration constants `PDFAComplianceModeEnum` for PDF/A-3a and PDF/A-3u formats correspondingly .

ENHANCED JBIG2 LOSSLESS COMPRESSION

The technology JBIG2 lossless compression format for black and white images was significantly improved in this release.

NEW PROPERTY IENGINE::VERSION

IEngine::Version returns the build number of the ABBYY FineReader Engine version you are using.

VERSION FOR FRENGINE.JNI.DLL WAS ADDED

From this release FREngine.jni.dll version was added. It coincides with the version of FRE you are using.

SCREENSHOT DETECTION

From now on our technologies are able to identify screenshots in documents. The document analyzer detects screenshots as image blocks if EnableTextExtractionMode property set to false (it is default value). Otherwise, text blocks can be detected on the screenshots. (helpdesk request 343728)

LAUNCH OF AINFO UTILITY FROM THE COMMAND LINE

It is possible now to call AInfo utility that is used to gather information for technical support from the command line in silent mode. (helpdesk request 388454)

IMPROVED KEY VALUE TABLES DETECTION

Significant improvements were made for key value tables' detection in this release.

SERBIAN CYRILLIC ICR SUPPORT

From now on Serbian Cyrillic is supported. Earlier It was not included in the distributive due to internal limitations.

IMPROVED MEMORY ALLOCATION

The developers have successfully managed to significantly reduce memory fragmentation. Thus, the error OUT_OF_MEMORY was fixed in many cases.

FIXED BUGS

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

The following terminology is used for frequency in the list:

General bugs are frequently reproduced.

Rare bugs are seldom reproduced on certain conditions.

Specific bugs are reproduced in custom images in very specific scenarios.

Frequency	Description	Subsystem	HD #	Office
General	The Arabic symbols ٣ and ٣ were often mixed up. The Classifiers for these characters were updated improved.	Recognizer	397942	RU
	Temporary files were recorded in the folder %ALLUSERSPROFILE%\Application Data\ABBY\SDK\11\Licenses and were not deleted. Now all temporary files are deleted after they become unnecessary.	Protection	391063	3A
	As the INSTALLDIR option is marked as required it shouldn't have default values. From now on INSTALLDIR option has no default values and users have to specify the path to the installation directory.	Install	401453	EU
	Internal Program Error after two successive runtime installations.	Install	394275	EU
	Problem with ASCII RTL. The output file with recognition results was empty.	API	393613	3A
	There was a change in the section of Help file: after the installation the files of fonts for synthesis and export to PDF are placed to Data folder not in Bin\Bin64.	Help	393406	US
	Table blocks were detected in the mode IPageAnalysisParams::DetectTables=FALSE. Issue was fixed; from now on the parameter will work as documented.	DA	389740	US
	In Java wrapper in ITableBlock interface the method FindBaseCellFromPoint wasn't supported. In Help file this limitation wasn't mentioned. From now on this method will be kept in the API as documented.	API	385844	EU
	The extra spaces were inserted in numbers with decimal point.	Recognizer	302474	US
	DataMatrix barcodes weren't detected	BarcodesDetector	383444	US
	There was problem with this Windows Installer package on Chinese Windows 7	Install	386824	US
	IRTFExportParams::KeepPages property didn't work correctly. There was a disparity between the API and the Help file. Now this property is	API		

	working as described in the Help file: the value of this property is used only if the PageSynthesisMode property is set to PSM_RTFFormatParagraphs or PSM_RTFFormatPlainText, otherwise it is ignored.			
Rare	There was problem with this Windows Installer package on Chinese Windows 7	Install	389818	3A
	An error occurs during the opening of pdf files using AddImageFileFromMemory Method with MultiProcessingMode = MPM_Auto	API	394159	UA
	Aztec code wasn't recognized.	BarcodesDetector	385839	EU
	Access violation in java project during classification.	API	361888	RU
Specific	The table was detected incorrectly.	DA	362040	RU
	The layout of output docx file is not the same with original pdf file.	Export	344028	US
	Some tables weren't detected on the specific image.	DA	359925	US
	IPE in the IFRDocument::Analyze method.	DA	341072	EU
	The QR barcodes weren't detected on the specific images.	BarcodesDetector	386532	EU
	After the export from a specific tif file to TXT and one of the columns wasn't displayed correctly.	Export	387248	EU
	An IPE occurs in the process of analyze of Korean document.	DA	381549	EU
	The result of export to TXT with RetainLayout = true was incorrect.	Export	350589	US
	The table wasn't detected on the specific picture.	Image	384343	RU
	The table wasn't detected on the specific picture.	DA	347439	US
The precision on Regular Expressions section about the fact that backslashes in the path must be doubled.	Help	372998	UA	

	IPE in the parallel processing mode.	DA	384351	EU
	The text in the specific pdf file wasn't detected completely.	DA	385218	EU
	The color was detected incorrectly on the specific image.	DA	344028	US
	The barcode PDF417 wasn't recognized.	BarcodesDetector	386257	US
	The table blocks weren't detected correctly on the specific image.	DA	389219	3A
	The numbers were detected as a picture with the property EnableTextExtractionMode = true on the specific image.	DA	391601	US
	The specific PDF files were opened very slowly with IFRDocument::AddImageFile method	Image/PDF	395329	EU
	The left column of the table wasn't detected.	DA	393181	EU
	The text wasn't recognized on the specific image.	DA	390245	US
	The Croatian characters Č, Ć, Ž were recognized as C, C, Z.	Synthesis	397849	US

KNOWN ISSUES AND WORKAROUNDS

SOME API IS NOT IMPLEMENTED

The following API is not implemented in FRE 10 and FRE 11:

- IFootnoteSeries::IsNumberingWithSuperscript. Always returns "false".
- IFootnoteSeries::PositionOnPage. Always returns "FPPT_SingleColumnSection".
- IFootnoteSeries::PositionInDocument. Always returns "FPDT_PageEnd".
- IFootnoteSeries::HasSeparator. Always returns "true".
- ITextPicture::ColumnNumber. Always returns "0".
- ICharParams::IsWordStart. Always returns "false". It is true only for character parameters got through IWordRecognitionVariants interface.
- IIncut::TextWrapping. Always returns "TW_Undefined".
- IRunningTitlesSeriesText::HasSeparator. Always returns "false".

The implementation is not planned.

ONPAGEPROCESSED COMES ONCE PER A DOCUMENT

During opening, synthesis, exporting stages the callback OnPageProcessed arrives only once per a document rather than after each page.

SKEW CORRECTION WORKS ON THE IMAGE WITHOUT TEXT

The skew is corrected on the image without text. The settings for skew correction are default. (helpdesk request 402457)

The fix of this issue is postponed.

IN PARALLEL MODE IWORDS COLLECTION IS EMPTY AFTER PROCESSING

After recognition with IEngine::ProcessPage in parallel mode IWords collection is empty.

SKEW IS NOT CORRECTED PROPERLY ON THE SPECIFIC IMAGES

Problem with text extraction accuracy and block detection due to inexact skew correction. (helpdesk request 372736)

PROBLEM WITH NUMBER RECOGNITION

The recognition results for numbers on the specific client's images are worse than the results in FRE 8. (helpdesk request 369894)

INCORRECT POSITIONING OF IMAGES IN THE RESULTS OF EXPORT TO HTML

Images from the PDF file are shifted in the HTML export results. (helpdesk request 344900)

THE RESULTS OF EXPORT TO XML IN SEVERAL CASES DON'T PASS XML VALIDATION

The results of export to XML with property WriteParagraphStyles = true for right to left text don't pass XML validation. The following error occurs: "The 'rtl' attribute is not declared"

THE TEXT DOESN'T FIT THE SIZE OF CELLS DURING EXPORT TO XLSX.

The text doesn't fit the cells when it is exported to XLSX in ExactCopy mode.

WRONG DETECTION OF FOOTERS AND HEADERS

In some cases extra footers and headers are detected on the processed documents.

INJECTTEXTLAYER METHOD WORKS INCORRECTLY ON SOME JAPANESE IMAGES

New IFREngine::InjectTextLayer methods that processes the input PDF file and creates and injects the text layer created from the recognized text in a searchable PDF file has a shortcoming. On some pictures with Japanese text the injected text layer contains wrong characters.

PDF/A VALIDATION REPORT

The following issues are known for PDF/A files produced by this release of FRE 11:

1. Adobe Acrobat 11.0.5 (Preflight 11.0.4) detects an error for PDF/A with attachments (PCM_Pdfa_3a):

“Embedded file does not have AF entity”

R2 GM PATCH –

PART NUMBER, BUILD NUMBER

Part#	1041/21
Build#	11.1.5.110

FIXED BUGS

This section contains a list of bugs that have been fixed in this release.

The following terminology is used for severity in the list:

Critical: the software will not run

High: unexpected fatal errors (includes crashes and data corruption)

Medium: a feature is malfunctioning

Low: a cosmetic issue

Severity	Description	Subsystem	HD #	Office
Critical	During the first setup of FREngine if the folder FREngineDataFolder (the parameter of GetEngineObjectEx function) didn't exist, FREngine crashed on the GetEngineObjectEx function with "Unhandled exception at 0x7769dcbb (ntdll.dll) in Reclamation.exe: 0xC0000374: A heap has been corrupted." The issue didn't reproduce during the second setup.	API	-	-
High	The project hung on the IFRDocument::Recognize method with the following settings: RecognizerParams::TextLanguage = English, CMC7 and RecognizerParams::TextTypes = (TT_MICR_CMC7 TT_Normal)	Recognizer	405817	US
Medium	Some text was lost during analysis stage on specific documents.	Analysis	407460	EU
Medium	Extra spaces appeared in the recognition results of addresses in envelopes in Italian.	Recognizer	364232	EU