**Recognition Server 4 Release 5**

**Release Notes**

|  |  |
| --- | --- |
| Author: | Helen Mamatova |
| Status: | Final |
| Last updated: | 11/27/2016 |
| Confidentiality level: | ABBYY |

**Table of Contents**

[Introduction 3](#_Toc465706551)

[About This Document 3](#_Toc465706552)

[About the Сurrent Release 3](#_Toc465706553)

[Technical Information 3](#_Toc465706554)

[Key Enhancements 3](#_Toc465706555)

[New Features and Improvements 4](#_Toc465706556)

[1. Improved processing of email messages 4](#_Toc465706557)

[1.1. Import of email messages in EML format 4](#_Toc465706558)

[1.2. Processing of attached email messages 4](#_Toc465706559)

[1.3. Properties of output files produced by converting email messages 5](#_Toc465706560)

[1.4. Email body processing improvements 5](#_Toc465706561)

[1.5. Ability to send output emails to recipients in the “To:” field of the input email 6](#_Toc465706562)

[2. Saving word confidence values to output Alto XML files 7](#_Toc465706563)

[3. Microsoft SharePoint 2016 support 7](#_Toc465706564)

[4. Microsoft Failover Cluster support 7](#_Toc465706565)

[5. ABBYY Recognition Server 4 IFilter improvements 7](#_Toc465706566)

[6. Known Issues 8](#_Toc465706567)

[About the Product 10](#_Toc465706568)

[About the Product Version 10](#_Toc465706569)

[Installing the Product Version 10](#_Toc465706570)

[License Usage 10](#_Toc465706571)

[History of Releases 10](#_Toc465706572)

[Release 4 for Symantec 10](#_Toc465706573)

[Release 4 10](#_Toc465706574)

[Release 3 with Japanese UI and Help 10](#_Toc465706575)

[Release 3 Patch 2 (for customer) 10](#_Toc465706576)

[Release 3 Patch 1 (for customer) 11](#_Toc465706577)

[Release 3 with Japanese UI 11](#_Toc465706578)

[Release 3 11](#_Toc465706579)

[Release 2 Patch 2 11](#_Toc465706580)

[Release 2 Patch 1 11](#_Toc465706581)

[Release 2 11](#_Toc465706582)

[Release 1 Multilingual 11](#_Toc465706583)

[Release 1 12](#_Toc465706584)

[Release 1 (specially for 3A) 12](#_Toc465706585)

[Arabic Edition 12](#_Toc465706586)

# Introduction

## About This Document

This document describes the features that have been implemented in ABBYY Recognition Server 4 Release 5.

## About the Сurrent Release

ABBYY Recognition Server 4 Release 5 provides improved processing of email content received from Exchange or POP3 servers, as well as a number of bug fixes. This document describes the new features in the release. The bug fixes are described in the separate document.

## Technical Information

*Part #:* ***1135/20****, build #* ***4.0.6.118****, OCR Technologies build #* ***13.0.28.117****, release date:* ***November 28, 2016.***

## Key Enhancements

* Improved processing of emails
* Support of Microsoft SharePoint 2016
* Microsoft Failover Cluster support
* Bug fixes

# New Features and Improvements

## Improved processing of email messages

### Import of email messages in EML format

Processing of email messages stored in EML format is now supported. EML messages can be imported within any workflow type: from shared folders, from SharePoint, from mailboxes (as files attached to email messages), etc.

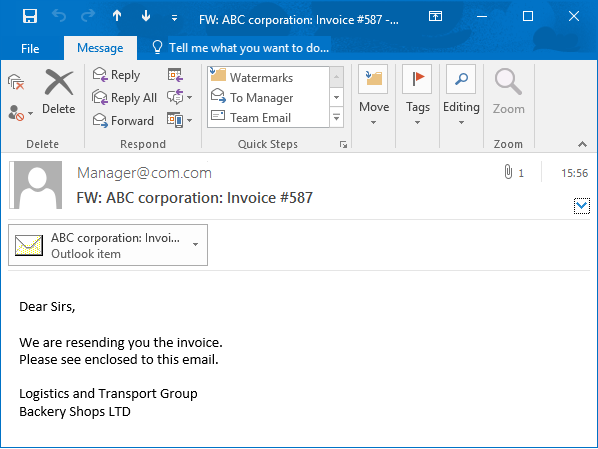
More information on configuring “Mail” workflows is provided below.

### Processing of attached email messages

Recognition Server allows processing email messages (including body and any attachments inside) attached to email messages imported from a POP3 or Microsoft Exchange mailbox.

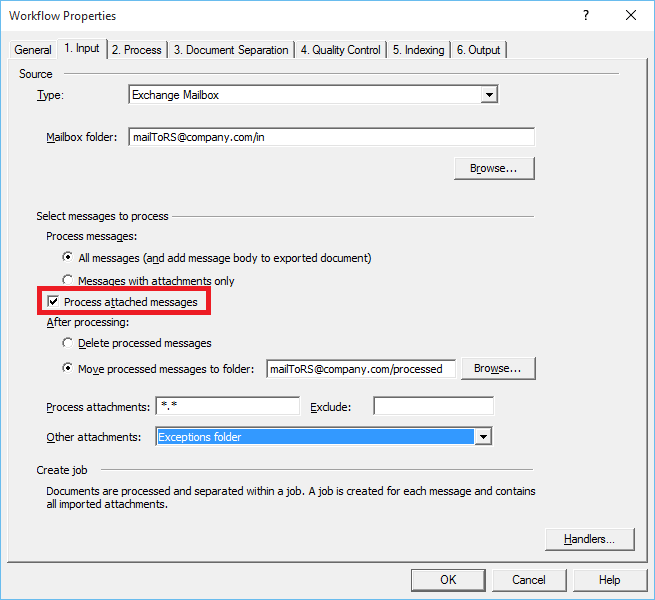
This may be useful in the following scenarios: case registration, customer relationship management, data investigation, and multi-channel input of incoming data when all correspondence contents are important or if previous emails are resent.

The following formats of attached messages are supported: Outlook email message, email of EML file type.



To enable the processing of both Outlook content messages and EML files, select the **Process attached messages** option.

If email messages in EML format should not be processed, this format should be added to the list of excluded formats. To achieve this, specify the **\*.eml** file extension in the **Exclude** field.



***Note:*** If the processing of attached email messages is enabled, all input settings will be applied to both main and attached email messages.

* Selecting the option **All messages (and add message body to exported document)** enables the processing of email messages both with and without attachments.
* If an attached message itself contains attached files or messages, the file extensions specified in the **Process attachments** and **Exclude** fields define which attached files or messages must be processed.

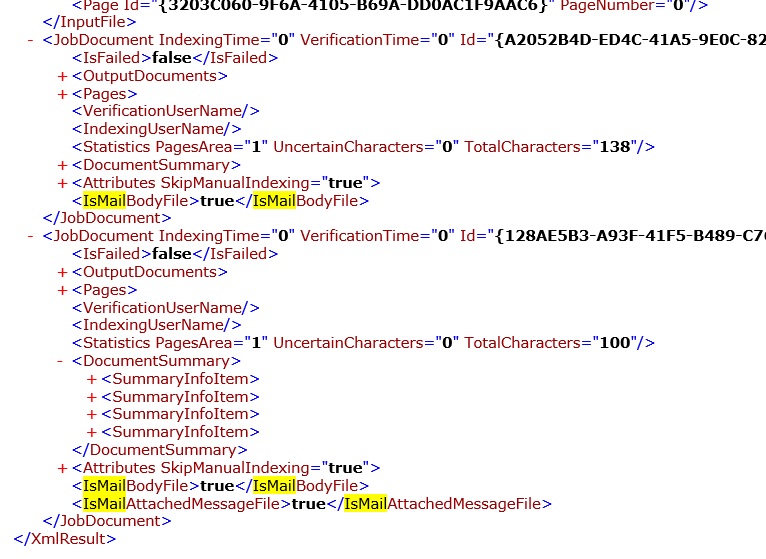
*Requested by: ABBYY EUHQ.*

### Properties of output files produced by converting email messages

When processing an email message and saving its body, attached files, and attached email messages into separate documents, it is useful to have information about the source of each output document. For example, it is useful to know whether a document was obtained from the body of a main message (e.g. request text) or from the body of an attached message (e.g. supplementary information). This information can be used to route documents to desired locations of the back-end system.

The following output document properties are available in the XML result file:

* IsMailBodyFile *// Shows whether the file was created from a message body (=True) or from an attachment (=False)*
* IsMailAttachedMessageFile *// Shows whether the file was created from an attached message (=True) or from a main message (=False)*



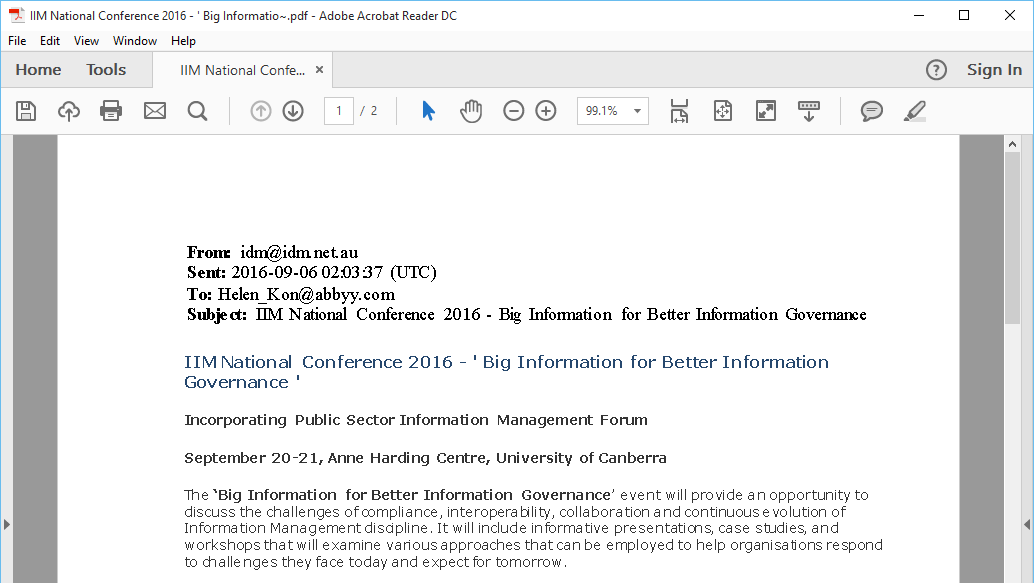
*Requested by: ABBYY EUHQ.*

### Email body processing improvements

Email contents processing has been improved:

1. An email header is now added to email body text.

Output files (or document pages) with email body now contain an email header with email date, sender, recipient, and subject. The language of the field names in the header depends on the Recognition Server interface language. The date and time in the **Sent** field of the email header are converted to UTC.



1. Email subjects are used as names for output files created from email bodies.

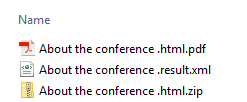
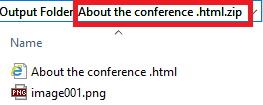
Output files produced by converting email bodies are now named using the text in their subject fields. If a subject field is empty, the output file name is **Email Message.<ext>.**

For emails with subjects longer than 50 characters, the filename will be truncated and the excess characters will be replaced with ~.

For example, the subject *IIM National Conference 2016 -' Big Information for Better Information Governance* will result in a file named *IIM National Conference 2016 -' Big Informatio~.pdf.*

1. Source emails are saved as HTML files or ZIP archives.

If the workflow output parameters require that source files should be saved into a folder, imported emails will be saved as HTML files. If an email contains pictures in its body, an HTML file with PNG picture files will be created and packed into a ZIP archive.

### Ability to send output emails to recipients in the “To:” field of the input email

Email conversion service for employees has been improved by adding the ability to send output emails to recipients in the “To:” field of the input email. This is useful when the user scans a document using an MFP and then sends an email with the scanned file to himself/herself (the user’s address is in the “To:” field).

To enable this feature:

* Select the **Reply to all** option in the workflow properties;
* Modify the *Configuration.xml* file as follows:
  + Set the ReplyToAddressesInToList parameter to true;
  + Specify excluded addresses (ExcludedAddresses parameter) if necessary.

To avoid a situation when the resulting email is also sent to the mail address used as an input source for Recognition Server import, it is strongly recommended that you add the Recognition Server address to the list of excluded addresses.

For example:

<ExportFormat ... OutputFlowType="Smtp" EmailSelectionMethod="ReplyToSender" ReplyToAddressesInToList="true" ...>

...

<ExcludedAddresses>RS\_OCR@corp.com</ExcludedAddresses>

...

</ExportFormat>

*Requested by: ABBYY EUHQ.*

## Saving word confidence values to output Alto XML files

An additional property, WC, which the contains word recognition confidence value, can now be exported into the output Alto XML file.

To enable this feature, modify the *Configuration.xml* file, setting the WriteWordConfidence parameter to True.

*Requested by: ABBYY EUHQ.*

## Microsoft SharePoint 2016 support

Microsoft SharePoint 2016 is now supported.

Export to Microsoft SharePoint libraries, library crawling, saving documents with index attributes, and other scenarios are available in both on-premise and online versions of Microsoft SharePoint 2016.

*Requested by: ABBYY 3A.*

## Microsoft Failover Cluster support

ABBYY Recognition Server can be deployed in a Microsoft Failover Cluster to ensure automatic switching to a functioning server and to reduce down times.

To implement this functionality, additional improvements have been made. It is now possible to store the server configuration file and temporary files of running tasks in a shared folder. This folder is available to any node with the Recognition Server installed and allows quick switching to another server for uninterrupted processing.

Instructions for installing and setting up Recognition Server in a Microsoft Failover Cluster are available in the Administrator’s Guide (currently available in English and Russian only).

*Requested by: ABBYY EUHQ, 3A, NAHQ.*

## ABBYY Recognition Server 4 IFilter improvements

ABBYY Recognition Server 4 IFilter is optimized for better performance. The timeouts of communication between IFilter components have been fine-tuned and logging of IFilter events has been disabled.

When tracking the IFilter operation and consulting with ABBYY Technical Support team, it is necessary to enable logging for IFilter.

To enable logging:

1. Create a folder for storing the log files. For example, *C:\RS\_Logs.*
2. Enable the IFilter frond-end logging in the windows registry:

[HKEY\_LOCAL\_MACHINE\SOFTWARE\Wow6432Node\ABBYY\RecognitionServer\4.0\Log\App]

"IsEnabled"="true"

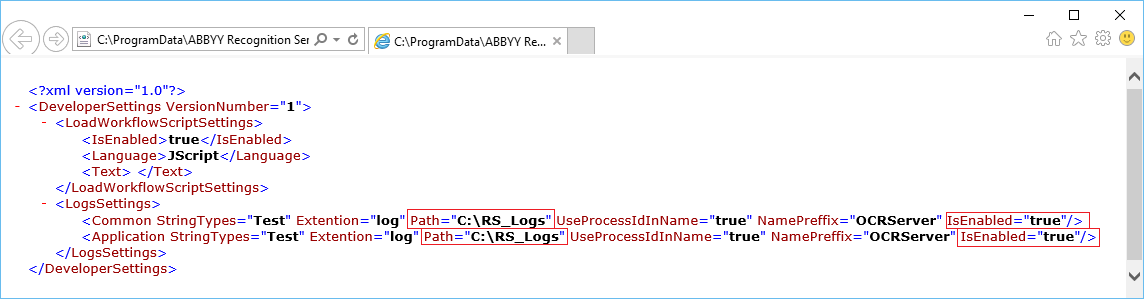
"FilePath"="C:\\RS\_Logs"

"IsEventLogEnabled"="false"

For 32-bit versions of Windows, Wow6432Node part of the path should be omitted.

1. (optionally) Enable the Recognition Server logging by modifying the *DeveloperSettings.xml* file:
   1. Edit the *C:\ProgramData\ABBYY Recognition Server 4.0\DeveloperSettings.xml* file
   2. In the Path parameters, specify the path to the folder, where logs will be kept
   3. Set IsEnabled parameter to the True

For example:



1. Send the log files from the folder to ABBYY Technical Support team.

*Requested by: ABBYY RU.*

## Known Issues

New known issues for ABBYY Recognition Server 4 are described in the table below. The full list of known issues is available on [MSK](http://msk/products/datacapture/Lists/Knowlege%20Base/Actual%20for%20RS4.aspx?View=%7B02218D54%2D51BF%2D4BD1%2D9FE9%2DA4E5249426BC%7D&ShowInGrid=HTML&InitialTabId=Ribbon%2EList&VisibilityContext=WSSTabPersistence).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Category** | **Titlе** | **Description** | **Solution or workaround** | **Reclamation Number** | **Release with issue** |
| Office Files Conversion | OpenOffice. Conversion fails, if a file can not be opened by OpenOffice | If a particular file can not be opened by OpenOffice application suite, then the conversion of such file also fails in Recognition Server with the default convertor based on Apache OpenOffice.  The following error message is shown in the logs:  *File 11.0.3.1770\_InvalidAuthors.xlsx Digital document conversion failed at the station Default Processing Station. Error: DDCModule: file:///C:/ProgramData/ABBYY Recognition Server 4.0/RS4WF/Images/DropDown/2150/j23\_11.0.3.1770\_InvalidAuthors.xlsx document could not be loaded..* | If Microsoft Office or LibreOffice can correctly open such files, use them for conversion in Recognition Server. | [835032](https://reclamationmanager.abbyy.com/view/835032) | Release 4; Release 5 |
| Office Files Conversion | Conversion via MS Excel fails, if the default printer is not available | ​Conversion of files via MS Excel convertor fails, if the default printer is not available on the machine with MS Office application installed. This may also happen due to using One Note as a default printer.  Unavailable printer/One Note printer cause the hanging of Print Spooler services (spoolsv.exe / splwow64.exe). | 1. It is recommended to delete One Note printer on the machine with Recognition Server processing station.  2. Add any local printer, that will be always available on the machine. This can be not necessarily a hardware device, but any virtual printer, for instance PDF printer.  3. Restart Print Spooler services (spoolsv.exe / splwow64.exe). | [832470](https://reclamationmanager.abbyy.com/view/832470) | Release 3; Release 4; Release 5 |
| Office Files Conversion | Conversion via MS Office under Local System account requires domain user credentials in workflow settings | ​Conversion of documents via Microsoft Excel may fail, if ABBYY Recognition Server 4 Processing Station service is running under Local System account. | The settings to perform conversion via MS Office applications depend on the MS Office version. **For MS Office 2010, 2013:**  - if the processing station is running under the Local System account, specify the domain user credentials (login/password) in the input event handler properties ("Settings for Input Office Files");  - if the processing station is running under the particular user account, there is no need in user credentials (login/password) in the handler properties, leave them unspecified. **For MS Office 2016:**  - it can be used only if the processing station is running under the Local System account and the domain user credentials are specified in the handler properties.  Start ABBYY Recognition Server 4 Processing Station service under the required account and configure the workflow settings, taking into consideration the information above. | [834937](http://rm/ViewReclamation.aspx?Id=834937) | Release 5 |
| Office Files Conversion | MS Office 2016 converts files only by station under Local System account | To use MS Office 2016 for conversion of office formats, it is necessary to run ABBYY Recognition Server 4 Processing Station service under Local System account and specify the domain user credentials in the handler properties of the workflow. |  | [834937](http://rm/ViewReclamation.aspx?Id=834937) | Release 5 |
| Server operation | Maximum 16 cores are shown in properties of Processing Station, if HyperThreading is enabled | If HyperThreading is used on a machine where ABBYY Recognition Server Procesing Station installed, then maximum 16 cores are shown in the properties of this station in the Administration console.  This happens due to the fact that only physical cores can be displayed in the properties (it doesn't depend on HyperThreading).  Recognition Server Processing Station can use maximum 32 cores for processing, it means that:   * without HyperThreading, 32 physical cores are used by a Processing Station * with HyperThreading, 32 virtual cores (16 physical cores) are used by a Processing Station   If the machine has more than 32 physical/32 virtual cores, then free cores can be used by the other Recognition Server components (server, API clients, etc) installed on this machine. |  | [826191](https://reclamationmanager.abbyy.com/view/826191)  HelpDesk:  [563616](https://helpdesk.abbyy.com/Request/Edit/563616) | Release 1; Release 1 (Special for 3A); Release 1 ML; Release 2; Release 3; Release 4; Release 5 |

# About the Product

## About the Product Version

ABBYY Recognition Server 4 provides new technology, including significantly improved recognition of texts in Arabic, new export settings, and other technology improvements. The main server features, such as stability, performance, and auto-recovery have been revised and improved. The new version can also process document libraries stored in read-only or editable folders. Other improvements include advanced logging, GUI changes, and bug fixes. The main changes are described in the history below.

## Installing the Product Version

Recognition Server 4 can be installed on the same computer where Recognition Server 3.5 or earlier is installed.

The configuration settings used in a previous version of ABBYY Recognition Server can be imported into ABBYY Recognition Server 4. For further information, please see the section “Upgrade from the previous versions of ABBYY Recognition Server” in the *System Administrator’s Guide*.

**Note**: please be aware that some changes have been made to the XML result file comparing to version 3.5. This may require changes in the software used for integrating ABBYY Recognition Server with data storage systems. For details, refer to the “XML Result” section in the Help file.

## License Usage

Recognition Server 4 cannot work with licenses generated for previous versions of Recognition Server (3.5 and earlier).

Recognition Server 4 can work with a license generated for Recognition Server Arabic Edition, but since some of the license parameters have been changed (ISIS option has been added), we recommend generating new licenses for Recognition Server 4 Release 1 (for 3A), Recognition Server Release 1 and other maintenance releases.

## History of Releases

### Release 4 for Symantec

*Part #:* ***1135/18****, build #* ***4.0.5.8891****, OCR Technologies build #* ***13.0.24.96****, release date:* ***September 01, 2016***

* Support of custom license parameter “Symantec DLP Connector”

### Release 4

*Part #:* ***1135/14****, build #* ***4.0.5.5022****, OCR Technologies build #* ***13.0.24.96****, release date:* ***February 02, 2016***

* Support of Microsoft SharePoint Online (Office 365)
* SharePoint libraries processing improvements
* Built-in component for conversion of digitally created documents

### Release 3 with Japanese UI and Help

*Part #:* ***1135/13****, build #* ***4.0.4.1447****, OCR Technologies build #* ***13.0.20.56****, release date:* ***October 9, 2015.***

* Translation of operator stations UI and Help to Japanese
* Bug fix for ABBYY USA Administration Console UI was translated into Japanese.

### Release 3 Patch 2 (for customer)

*Part #:* ***1135/12****, build #* ***4.0.4.1438****, OCR Technologies build #* ***13.0.20.56****, release date:* ***September 22, 2015.***

* Bug fix. Processing Stations can now be connected to the Server, if TCP/IP protocol is used as a protocol for interaction between Recognition Server components. The “Access is denied” bug has been fixed.

### Release 3 Patch 1 (for customer)

*Part #:* ***1135/11****, build #* ***4.0.4.1437****, OCR Technologies build #* ***13.0.20.56****, release date:* ***September 07, 2015.***

* Bug fix. Original file can now be overwritten with the newly generated file with the same file name and extension, when processing documents in Microsoft SharePoint library (e.g. to avoid the duplication of PDF files).

### Release 3 with Japanese UI

*Part #:* ***1135/10****, build #* ***4.0.4.1434****, OCR Technologies build #* ***13.0.20.56****, release date:* ***July 17, 2015.***

* Administration Console UI was translated into Japanese.

### Release 3

*Part #:* ***1135/9****, build #* ***4.0.4.1425****, OCR Technologies build #* ***13.0.20.54****, release date:* ***June 15, 2015.***

* Conversion of Office file formats
* Processing the entire SharePoint portal with child sites within one workflow
* Saving output files in input folders
* Writing original documents as attachments to PDF/A and PDF documents
* Improvements in export to ALTO XML
* Sending notifications to the Administrator via an SMTP server

### Release 2 Patch 2

*Part #:* ***1135/8****, build #* ***4.0.3.1180****, OCR Technologies build #* ***13.0.15.138****, release date:* ***February 6, 2015.***

* Improved work with multiple workflows (more than 16) and indexing of documents that contain many index fields.

### Release 2 Patch 1

*Part #:* ***1135/7****, build #* ***4.0.3.1175****, OCR Technologies build #* ***13.0.15.138****, release date:* ***January 16, 2015.***

* Possibility to specify the color of filling the document edges (“triangles”) after automatic deskew

*(see the feature* [*description*](#_Setup_the_color)*)*

### Release 2

*Part #:* ***1135/6****, build #* ***4.0.3.1167****, OCR Technologies build #* ***13.0.15.131****, release date:* ***November 14, 2014***

* [Improved MRC compression method](#_Improved_MRC_compression_1) (producing the compressed PDF files of the minimum size)
* [Using IFilter for processing PDF files in MS SharePoint](#_Using_IFilter_for)
* [Processing the SharePoint document libraries:](#_Workflow_settings)
  + [Crawling of the whole site (including multiple libraries and folders)](#_Processing_SharePoint_libraries)
  + [Periodical re-crawling settings](#_Periodical_crawling_of)
* [Export to specific column types in SharePoint](#_Export_to_specific) (support of Date, Number, and some other formats)
* [Export to PDF/A-3](#_Improved_MRC_compression)

### Release 1 Multilingual

*Part #: 1135/5, build # 4.0.2.952, OCR Technologies build number 13.0.13.21, release date: 14/08/2014*

* Translation of UI and help on the following languages:
  + French
  + German
  + Italian
  + Spanish
  + Chinese
  + Portuguese (Brazil)
  + Czech
  + Hungarian
  + Polish
* Bug fix for ABBYY USA

### Release 1

*Part #: 1135/4, build # 4.0.2.943, OCR Technologies build number 13.0.13.15, release date: 19/05/2014*

* Improved failure recovery
* Ability to limit the number of processed pages
* Verification and Indexing Station improvements
  + Selecting documents from a queue
  + Timeout settings
  + Saving changes on the stations
* Indexing Station improvements
  + Importing document types from an external source

### Release 1 (specially for 3A)

*Part #: 1135/3, build # 4.0.1.795, OCR Technologies build number 13.0.8.108, release date: 29/01/2014*

* Improved server operation
  + Redundancy
  + Reports and statistics
* Enhanced work with PDF files
* Ability to process documents in read-only folders
* Processing documents placed in SharePoint libraries
* Latest technology version

### Arabic Edition

*Part #: 1135/2, build # 4.0.0.461, OCR Technologies build number 13.0.0.58, release date: 06/05/2013*

* Improved recognition of Arabic texts
* Ability to process documents in read-only folders
* Improved logging
* Bug fixes