

ABBYY[®] FineReader

Version 11

User's Guide

Information in this document is subject to change without notice and does not bear any commitment on the part of ABBYY.

The software described in this document is supplied under a license agreement. The software may only be used or copied in strict accordance with the terms of the agreement. It is a breach of the "On legal protection of software and databases" law of the Russian Federation and of international law to copy the software onto any medium unless specifically allowed in the license agreement or nondisclosure agreements.

No part of this document may be reproduced or transmitted in any form or by any means, electronic or other, for any purpose, without the express written permission of ABBYY.

© ABBYY Production LLC, 2013. All rights reserved.

ABBYY, the ABBYY logo, ABBYY FineReader, ADRT are either registered trademarks or trademarks of ABBYY Software Ltd.

© 1984-2008 Adobe Systems Incorporated and its licensors. All rights reserved.

Protected by U.S. Patents 5,929,866; 5,943,063; 6,289,364; 6,563,502; 6,185,684; 6,205,549; 6,639,593; 7,213,269; 7,246,748; 7,272,628; 7,278,168; 7,343,551; 7,395,503; 7,389,200; 7,406,599; 6,754,382; Patents Pending.

Adobe® PDF Library is licensed from Adobe Systems Incorporated.

Adobe, Acrobat®, the Adobe logo, the Acrobat logo, the Adobe PDF logo and Adobe PDF Library are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

© 1996-2007 LizardTech, Inc. All rights reserved. DjVu® is protected by U.S. Patent No. 6.058.214. Foreign Patents Pending.

© 2008 Celartem, Inc. All rights reserved.

© 2011 Caminova, Inc. All rights reserved.

Powered by AT&T Labs Technology.

© 2011 University of New South Wales. All rights reserved.

Fonts Newton, Pragmatica, Courier © 2001 ParaType, Inc.

Font OCR-v-GOST © 2003 ParaType, Inc.

© 2002-2008 Intel Corporation.

© 2010 Microsoft Corporation. All rights reserved.

Microsoft, Outlook, Excel, PowerPoint, Windows Vista, Windows are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

© 1991-2008 Unicode, Inc. All rights reserved.

© 2010, Oracle and/or its affiliates. All rights reserved

OpenOffice.org, OpenOffice.org logo are trademarks or registered trademarks of Oracle and/or its affiliates.

JasPer License Version 2.0:

© 2001-2006 Michael David Adams

© 1999-2000 Image Power, Inc.

© 1999-2000 The University of British Columbia

© 1996-2011, Amazon.com, Inc. or its affiliates. All rights reserved.

KINDLE®, KINDLE logo and other Amazon.com graphics, button icons are registered trademarks, trademarks or trade dress of Amazon in the U.S. and/or other countries...

EPUB®, is a registered trademark of the IDPF (International Digital Publishing Forum)

iPad is trademarks of Apple Inc., registered in the U.S. and other countries.

All other trademarks are the property of their respective owners.

Contents

Introducing ABBYY FineReader	4
What is ABBYY FineReader	4
What's New in ABBYY FineReader 11	5
The ABBYY FineReader 11 Interface	6
The Main Window	6
Toolbars	8
Customizing the ABBYY FineReader Workspace	9
The Options Dialog Box	9
Working with ABBYY FineReader	12
ABBYY FineReader Tasks	13
Custom Automated Tasks	18
ABBYY FineReader Step-by-Step	21
ABBYY FineReader Document	26
Taking Into Account Some of the Features of Your Paper Document	29
Image Acquisition Tips	32
Tips for Improving OCR Quality	42
Checking and Editing the Recognized Text	51
Working with Complex-Script Languages	56
Saving the Results	59
Advanced Features	75
Working in Other Applications	75
Using Area Templates	76
Recognition with Training	76
User Languages and Language Groups	79
Group Work in a LAN	82
ABBYY Hot Folder	83
Appendix	89
Glossary	89
Supported Image Formats	93
Supported Saving Formats	95
Languages and Their Fonts	96
Regular Expressions	98
Shortcut Keys	100
How to Buy an ABBYY Product	106
About ABBYY	106
Activating and Registering ABBYY FineReader	107
ABBYY FineReader Activation	107
ABBYY FineReader Registration	108
Privacy Policy	109
Technical Support	110

Introducing ABBYY FineReader

This chapter provides an overview of ABBYY FineReader and its features.

Chapter Contents

- What Is ABBYY FineReader
- What's New in ABBYY FineReader

What is ABBYY FineReader

ABBYY FineReader is an optical character recognition (OCR) system. It is used to convert scanned documents, PDF documents, and image files, including digital photos, into editable formats.

ABBYY FineReader advantages

Fast and accurate recognition

- The OCR technology used in ABBYY FineReader lets users quickly and accurately recognize and retain the original formatting of any document (including text on background images, colored text on colored backgrounds, text wrapped around an image, etc.).
- Thanks to ABBYY's adaptive document recognition technology (**ADRT®**), ABBYY FineReader can analyze and process a document as a whole, instead of page by page. This approach retains the source document's structure, including formatting, hyperlinks, e-mail addresses, headers and footers, image and table captions, page numbers, and footnotes.
- ABBYY FineReader can recognize documents written in one or several of **189 languages**, including Arabic, Vietnamese, Korean, Chinese, Japanese, Thai, and Hebrew. ABBYY FineReader also features auto-detection of document languages.
- ABBYY FineReader is largely immune to printing defects and can recognize texts printed in virtually any font.
- The program also includes a wide range of options for output data: documents can be saved in a variety of formats, sent by e-mail, or transferred to other applications for further processing.

Ease of use

- ABBYY FineReader's user-friendly and intuitive results-driven interface allows you to use the program without any additional training. New users can master the main functions in almost no time at all. Users can change the interface language directly in the program.
- **ABBYY FineReader Tasks** provide instant access to the most frequently used conversion tasks, making it possible to convert scanned documents, PDFs, and image files into editable formats with just one click.
- Seamless integration with Microsoft Office and Windows Explorer lets you recognize documents directly from Microsoft Outlook, Microsoft Word, Microsoft Excel, and Windows Explorer.
- FineReader also features a Help menu with examples that illustrate how to use the program to accomplish a number of complex conversion-related tasks.

What's New in ABBYY FineReader 11

Increased processing efficiency

With the new Black-and-White mode, FineReader 11 provides 30% faster processing when you don't need color. Furthermore, the program effectively takes advantage of multi-core processors for even faster conversion.

Flexible e-book creation

Scan paper books and convert them to EPUB and FB2 formats to read them on-the-go on your iPad, tablet or preferred portable device. Or send them directly to your Kindle account. Convert paper books or articles to appropriate e-book formats to add them to your electronic library or archive.

Native support for OpenOffice.org Writer

FineReader 11 recognizes and converts images of documents and PDF files directly to OpenOffice.org Writer format (**ODT**), accurately preserving their native layout and formatting. Now you can easily add documents to your *.odt archive with a few mouse clicks.

Enhanced user interface

- Enhanced style editor allows you to set up all the style parameters in one friendly window. All the changes are applied across the entire document at once.
- Sort pages into different FineReader documents for better layout retention.
- Easier access to predefined and custom conversion tasks lets you start converting documents as soon as you launch the program.

Next generation Camera OCR with extended set of powerful image editing tools

FineReader 11 offers a wide range of new powerful image editing tools, including brightness and contrast sliders and level tools enabling you to get more accurate results by improving image quality.

Increased OCR accuracy and layout retention with improved ADRT 2.0

Minimize the time required to edit converted documents thanks to better detection of document styles, text in the margins, headers and footers.

Optimal PDF output

The three pre-defined image settings for PDF files provide optimal results depending upon your needs – best quality, compact size or balanced mode.

New recognition languages* – Arabic, Vietnamese, and Turkmen (Latin alphabet)

Quickly convert paper business cards into electronic contacts with Business Card Reader (available only in the Corporate Edition).

*Different sets of recognition languages may be available in different editions of the product.

The ABBYY FineReader 11 Interface

ABBYY FineReader's user-friendly, intuitive, and results-driven interface makes it easy to use the program without any additional training. It takes new users almost no time at all to master the main functions.

ABBYY FineReader's interface can be customized.

You can customize your workspace to suit your individual needs.

- Change the window position and size
- Set up the the quick access bar for access to the most frequently used commands
- Set up the keyboard shortcuts — you can change the preset key combinations and add your own to run specific program commands
- Select your preferred interface language, etc.

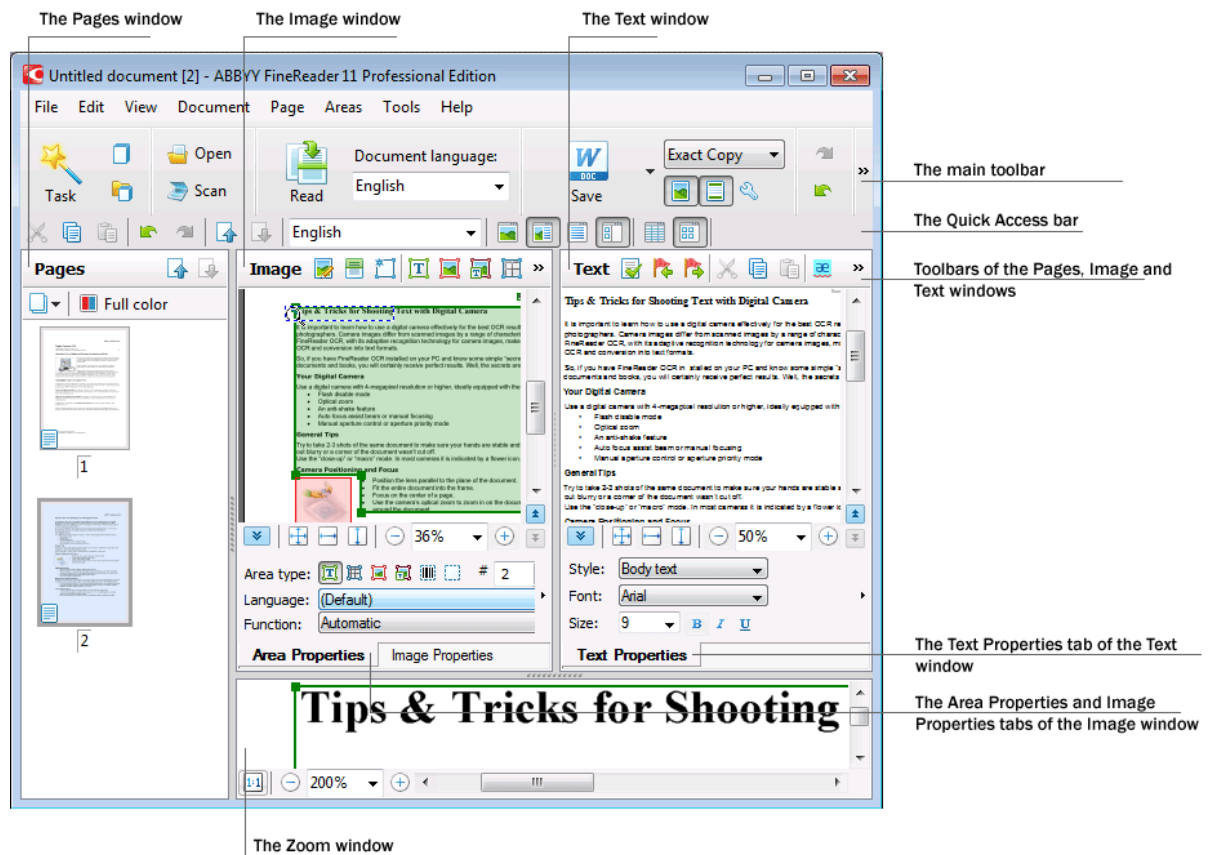
Chapter Contents

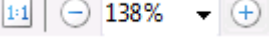
- The Main Window
- Toolbars
- Customizing the ABBYY FineReader Workspace
- The Options Dialog Box

The Main Window

When you start ABBYY FineReader, its **main window** opens together with a **New Task** window where you can set up and launch a built-in quick task or a custom automated task.

The main window displays the ABBYY FineReader document that is currently open. For details, see ABBYY FineReader documents.



- In the **Pages** window, you can view the pages of the current FineReader document. There are two page viewing modes: icons (which shows a small image) or a list of document pages and their properties. You can toggle between these modes from this window's shortcut menu, from the **View** menu, and the **Options** dialog box (**Tools>Options...**).
- The **Image** window displays an image of the current page. You can edit image areas, page images, and text properties in this window.
- In the **Text** window, you can view recognized text. You can also check spelling and format and edit recognized text in the **Text** window.
- The **Zoom** window displays an enlarged image of the line or processed image area currently being edited. While the **Image** window displays the general page view, the **Zoom** window provides an easy way to view the image in greater detail, adjust the area type and position, or compare uncertain characters with the enlarged image. You can change the image scale in the **Zoom** window from the panel  in the bottom part of the window.

Tip. You can move images in the **Image** and **Zoom** windows with your mouse while holding down the space bar.

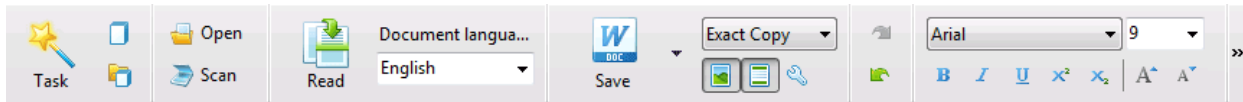
Window toggle keys

- Press CTRL+TAB to toggle between windows
- Press ALT+1 to activate the **Pages** window
- Press ALT+2 to activate the **Image** window
- Press ALT+3 to activate the **Text** window

Toolbars

The main window of ABBYY FineReader features several toolbars. Toolbar buttons provide easy access to program commands. The same commands can be executed from the program menu or with keyboard shortcuts.

The **Main Toolbar** has a fixed set of buttons for all basic actions: opening a document, scanning a page, opening an image, recognizing pages, saving results, etc.



The **Quick Access Bar** can be customized by adding special buttons for virtually any command in the main menu. By default, this bar is not displayed in the main window. To display it, select the **Quick Access Bar** under **View>Toolbars**, or from the main toolbar's shortcut menu.



In the **Pages**, **Image**, and **Text** windows, the toolbars are positioned at the top. The buttons on the toolbar in the **Pages** window cannot be changed. The toolbars in the **Image** and **Text** window are fully customizable.



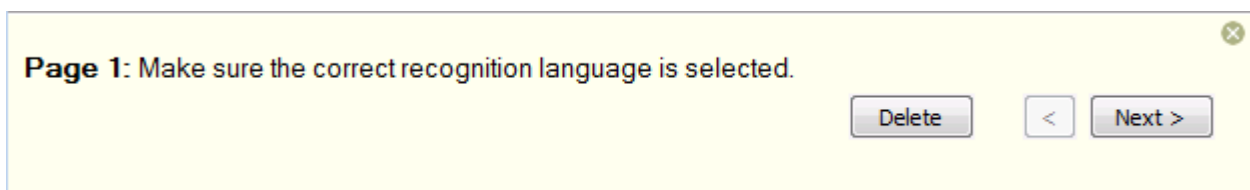
To customize a toolbar, i.e. to add or remove any of the buttons:

1. Open the **Customize Toolbars and Shortcuts** dialog box (**Tools>Customize...**).
2. On the **Toolbars** tab in the **Categories** field, select the desired category.


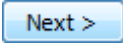
Note: When a different category is selected, the commands on the **Commands** list will change according to the set of commands on the main menu.

3. Select a command from the **Commands** list and click the **>>** (or **<<**) button.
The selected command will appear on (or disappear from) the toolbar's command list.
4. In order to save your changes, click the **OK** button.
5. In order to reset the Quick Access Bar to its default view, click the **Reset All** button.

The warnings and errors pane displays error warnings and pop-up messages when the program is running.





To open the warnings and errors pane, select **Show Warnings and Errors Pane** on the **View** menu or on the shortcut menu of the main toolbar.

Errors and warnings are displayed for the page that is currently selected in the **Pages** window. To scroll the warnings and errors, use the  and  buttons.

Customizing the ABBYY FineReader Workspace

Changing FineReader interface

- You can temporarily hide the windows you do not use. In order to hide/show windows, use the **View** menu options or keyboard shortcuts:
 - **F5** for the **Pages** window
 - **F6** for the **Image** window
 - **F7** for the **Image** and **Text** windows
 - **F8** for the **Text** window
 - **CTRL+F5** for the **Zoom** window
- You can change the window size by dragging the window spacers with your mouse.
- You can also change the position of the **Pages** and **Zoom** windows. Use the commands in the **View** menu or the window shortcut menu.
- You can customize the way warnings and error messages are displayed while the program is running. The **Show Warnings and Errors Pane** command in the **View** menu toggles the warnings and errors pane on and off.
- To show/hide the **Properties** panel in the **Image** or **Text** window, right-click anywhere in the window and then select **Properties** from the shortcut menu. Alternatively, you can click / at the bottom of these windows.
- Some window settings can be customized on the **View** tab of the **Options** dialog box.

Changing the interface language

The interface language is selected when ABBYY FineReader is installed. This language will be used for all the messages, dialog boxes, buttons, and menu items. To change the interface language, follow the instructions below:

1. Open the **Options** dialog box and select the **Advanced** tab (**Tools>Options...**).
2. Select the desired language from the **Interface language** drop-down menu.
3. Click the **OK** button.
4. Restart ABBYY FineReader.

The Options Dialog Box

In the **Options** dialog box, you can select document processing options and customize the appearance of the program.

To open the **Options** dialog box, select **Tools>Options...** or click the wrench icon on the toolbar at the top of the screen. Alternatively, you can right click the toolbar and select **Options...** This dialog box can also be accessed from the **Open Image** and **Save As** dialog boxes.

There are six tabs in the dialog box:

- **Document**

On this tab, you can select:

- Document language(s)
- Document print type
- Color mode
- Document properties

This tab also displays the path to the current ABBYY FineReader document.

- **Scan/Open**

On this tab, you can select automated image and document processing options:

- Automatically read acquired page images
- Automatically analyze acquired page images
- Enable image pre-processing
- Detect page orientation
- Split facing pages

You can also select your scanner driver and scanning interface.

- **Read**

On this tab, you can select recognition options:

- Thorough reading or fast reading
- Use built-in patterns
- Fonts to be used in recognized text
- Look for barcodes

Here you can also specify whether user patterns and languages should be saved to or loaded from a file.

- **Save**

On this tab, you can select the output format in which to save the recognized text:

- RTF/DOC/DOCX/ODT
For details, see Saving in RTF/DOC/DOCX/ODT.
- XLS/XLSX
For details, see Saving in XLS/XLSX.
- PDF
For details, see Saving in PDF.
- PDF/A
For details, see Saving in PDF/A.

- **HTML**
For details, see Saving in HTML.
- **PPTX**
For details, see Saving in PPTX.
- **TXT**
For details, see Saving in TXT.
- **CSV**
For details, see Saving in CSV.
- **E-book**
For details, see Saving E-Books.
- **DjVu**
For details, see DjVu.

- **View**

On this tab, you can:

- Customize the appearance of the **Pages** window, choosing between **Thumbnails** and **Details** views
- Customize the **Text** window, where you can have the unreliably recognized characters highlighted in a color of your choice, hide/show non-printed characters, and select the font to be used for plain text.
- Select the colors and width of frames that enclose the areas displayed in the **Image** window

- **Advanced**

On this tab, you can:

- Select verification options for unreliably recognized characters
- Specify whether spaces before and after punctuation marks should be corrected
- View and edit user dictionaries
- Specify whether ABBYY FineReader should automatically check for and install updates
- Select the language of the FineReader menus and messages
- Specify whether the last ABBYY FineReader document should be opened when you start the application
- Specify whether document warnings should be displayed in a separate pane
- Select if you wish to participate in ABBYY's software improvement program
- Reset the options to the program defaults

Working with ABBYY FineReader

This chapter provides an overview of ABBYY FineReader: how to quickly produce an e–document in a format convenient for editing and use, and how to perform complex and nonstandard text–recognition tasks.

Chapter Contents

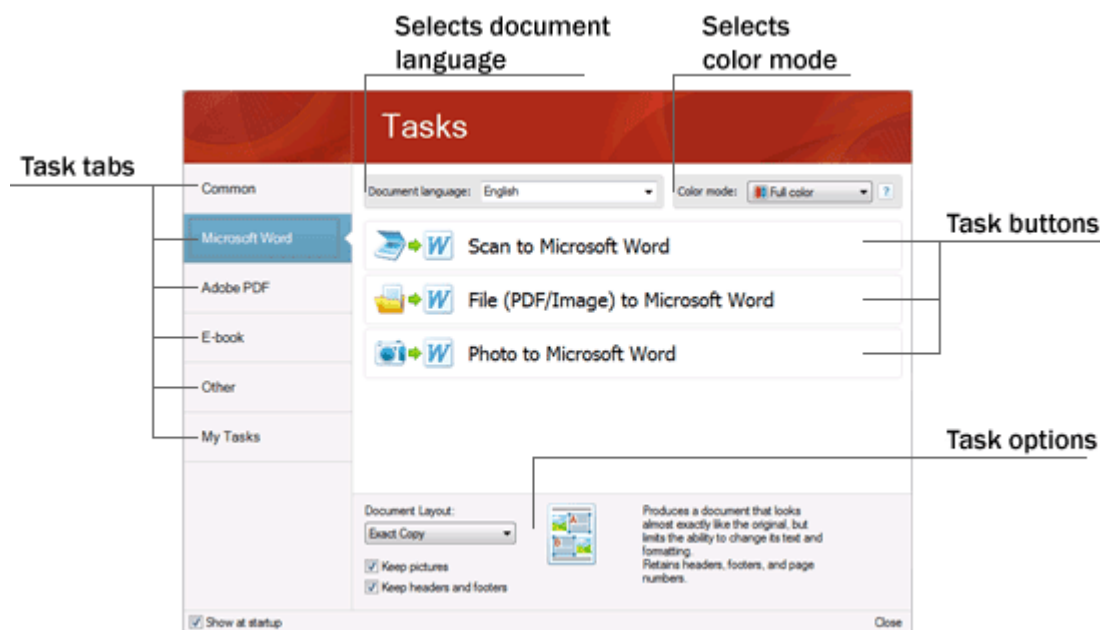
- ABBYY FineReader Tasks
 - Common Tasks
 - Microsoft Word Tasks
 - Adobe PDF Tasks
 - E–book Tasks
 - Other Quick Tasks
- Custom Automated Tasks
 - How to Create a New Automated Task
 - Task Steps
 - Managing Automated Tasks
- ABBYY FineReader Step–by–Step
 - Getting an Image
 - Recognition
 - Checking and Editing
 - Saving OCR Results
 - E–mailing OCR Results
 - Sending OCR Results to Kindle
- ABBYY FineReader Document
 - General
 - Working With ABBYY FineReader Documents
 - Splitting an ABBYY FineReader Document
 - Page Numbering in ABBYY FineReader Documents
- Taking Into Account Some of the Features of Your Paper Document
- Image Acquisition Tips
- Tips for Improving OCR Quality

- Checking and Editing the Recognized Text
- Working with Complex-Script Languages
- Saving the Results in different formats

ABBYY FineReader Tasks

Recognizing documents often involves repetitive tasks, such as scanning a document, recognizing it, and saving the results in a particular format. ABBYY FineReader's quick tasks enable you to save time by executing the most common tasks with a single click.

Quick tasks are launched from the **New Task** window, which opens by default after the application loads. If the **New Task** window is not displayed, click the **New Task** button on the main toolbar.



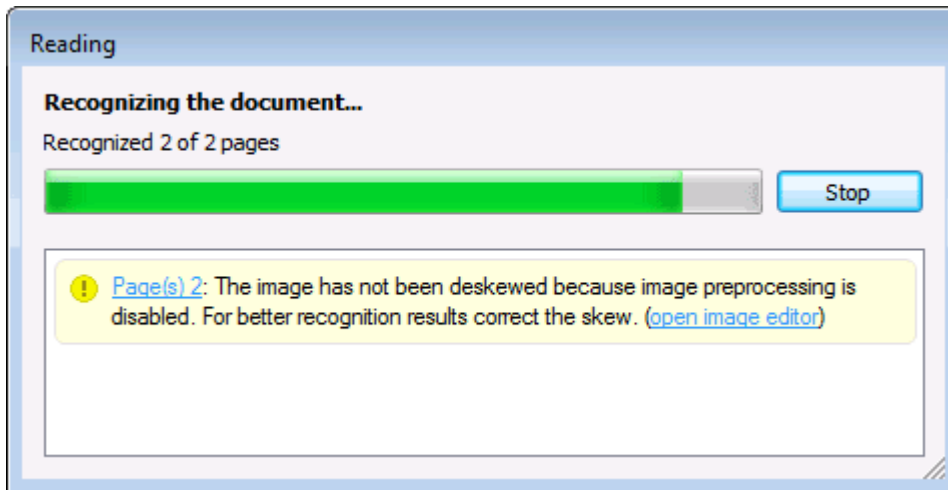
1. In the **New Task** window, select the desired task:
 - Common Tasks
 - Microsoft Word tasks
 - Adobe PDF tasks
 - E-book tasks
 - Other tasks
 - My Tasks
2. Select recognized languages from the **Document language** drop-down menu. Additional options are available (at the bottom of the window) for automated tasks used to convert data into a Microsoft Word document or an Adobe PDF document.
3. In the **Color mode** drop-down list, select full color or black-and-white.
 - **Full color** preserves the original image colors.
 - **Black-and-white** converts images to black-and-white, reducing the size of the ABBYY FineReader document and speeding up OCR.

Important! Once the document is converted to black-and-white, you will not be able to restore the colors. To get a color document, open the file with color images or scan the paper document in color mode.

4. Click the appropriate button for your automated task.

When you run a task it will use the options selected in the **Options** dialog box (**Tools>Options...**).

While an automated task is running, the screen shows the **task execution bar** with the task execution indicator, a list of automated task steps, tips, and warnings.



Note: In order to recognize documents with complex structures, you can customize and start each processing stage manually. For details, see ABBYY FineReader Step-by-Step.

You can also create custom automated tasks by selecting all the necessary steps. For details, see Managing Automated Tasks.

Common Tasks

Click the **Common** tab in the **New Task** window to access the most frequently used quick tasks.

1. Select the document languages from the **Document language** drop-down menu located in the upper section of the window.
2. In the **Color mode** drop-down list, select full color or black-and-white.
Important! Once the document is converted to black-and-white, you will not be able to restore the colors.
3. Click the appropriate button for your task:
 - **Scan to Microsoft Word** scans a paper document and converts it to a Microsoft Word document
 - **File (PDF/Image) to Microsoft Word** converts PDF documents and images files to a Microsoft Word document
 - **Scan and Save Image** scans a document and saves the resulting images. Once the scanning is finished, a dialog box opens prompting you to save the images.
 - **Scan to PDF** scans a paper document and converts it to an Adobe PDF document

- **Photo to Microsoft Word** converts a digital photo to a Microsoft Word document
 - **Scan** scans a paper document
 - **Open** opens a PDF document or image file
 - **Create Document...** creates a new ABBYY FineReader document
4. As a result, a new document with the recognized text will open in the application you have selected. When the **Scan and Save Image** task is completed, the image saving dialog box will open.

Important! Important! When a built-in task is launched, ABBYY FineReader uses the options selected at the moment of launch. If you adjust the options, launch the task once again to recognize the texts with these new options.

Microsoft Word Tasks

The tasks on the **Microsoft Word** tab of the **New Task** window will help you convert documents to a Microsoft Word document.

1. In the **Document language** drop-down list at the top of the window, select the document languages.
2. In the **Color mode** drop-down list, select full color or black-and-white.
Important! Once the document is converted to black-and-white, you will not be able to restore the colors.
3. In the bottom part of the window, select the desired conversion settings:
 - Document format saving mode
 - **Keep pictures** if you need to retain the images in the target document
 - **Keep headers and footers** if you need to retain the headers and footers in the target document
4. Click the appropriate button for your task:
 - **Scan to Microsoft Word** if you need to scan and convert a paper document
 - **File (PDF/Image) to Microsoft Word** if you need to convert PDF documents and image files
 - **Photo to Microsoft Word** if you need to convert digital photos
5. When the document processing is complete, a new Microsoft Word document containing the recognized text will open.
Note: If Microsoft Word is not installed on your computer, clear the **Open document after saving** option in the **Save As** dialog box.

Important! When you run a quick task, data are converted based on the program's current options. If you have changed application options, run the task again to recognize text with the new options.

Adobe PDF Tasks

The tasks on the **Adobe PDF** tab in the **New Task** window will help you convert various images to Adobe PDF format.

1. In the **Document language** drop-down list at the top part of the window, select the document languages.
2. In the **Color mode** drop-down list, select full color or black-and-white.
Important! Once the document is converted to black-and-white, you will not be able to restore the colors.
3. Select a document saving mode:
 - **Text and pictures only**
 This option saves only the recognized text and the associated pictures. The page will be fully searchable and the size of the PDF file will be small. The appearance of the resulting document may slightly differ from the original.
 - **Text over the page image**
 This option saves the background and pictures of the original document and places the recognized text over them. Usually, this PDF type requires more disk space than **Text and pictures only**. The resulting PDF document is fully searchable. In some cases, the appearance of the resulting document may slightly differ from the original.
 - **Text under the page image**
 This option saves the entire page image as a picture and places the recognized text underneath. Use this option to create a fully searchable document that looks virtually the same as the original.
 - **Page image only**
 This option saves the exact image of the page. This type of PDF document will be virtually indistinguishable from the original but the file will not be searchable.
4. From the **Picture** drop-down list, select the desired quality of the pictures in the document.
5. Select PDF or PDF/A.
6. Click the task you wish to use:
 - **Scan to PDF** scans a paper document and converts it to Adobe PDF
 - **File (Image) to PDF**, converts a PDF document or an image file to a searchable Adobe PDF document
 - **Photo to PDF**, converts a digital photo to an Adobe PDF document
7. The resulting PDF document will be opened in a PDF viewing application.
Note: If there is no PDF viewing application installed on your computer, clear the **Open document after saving** option in the **Save As** dialog box.

Important! When you run a quick task, data are converted based the program's current options. If you have changed application options, run the task again to recognize text with the new options.

Tip. When saving your scanned document to PDF, you can set passwords to protect your PDF document from unauthorized opening, printing, or editing. For details, see PDF security settings.

E-book Tasks

Use the tasks on the **E-book** tab of the **New Task** window to convert images to EPUB, FB2, and HTML, which are widely used for electronic books. The EPUB and FB2 formats are based on XML

and store each e-book in a single file. They are compatible with many e-book reading applications and devices.

1. Select the document languages from the **Document language** drop-down menu located in the upper section of the window.
2. In the **Color mode** drop-down list, select full color or black-and-white.
Important! Once the document is converted to black-and-white, you will not be able to restore the colors.
3. Select a format for your e-book: EPUB, FB2 or HTML.
4. Click the button with the name of the task that you wish to use:
 - **Scan to EPUB/FB2/HTML** scans a paper document and converts it to an e-book format of your choice
 - **File (PDF/Image) to EPUB/FB2/HTML** converts PDF documents and image files
 - **Photo to EPUB/FB2/HTML** converts digital photos
5. The document with the recognized text will be opened in the appropriate application.
Note: If there is no e-book reading application installed on your computer, clear the **Open document after saving** option in the **Save As** dialog box.

Important! When a built-in task is launched, ABBYY FineReader uses the options selected at the moment of launch. If you adjust the options, launch the task once again to recognize the texts with these new options.

Other Tasks

Use the **Other** tab in the **New Task** window to access the other quick tasks available in ABBYY FineReader.

1. In the **Document language** drop-down list at the top of the window, select the document languages.
2. In the **Color mode** drop-down list, select full color or black-and-white.
Important! Once the document is converted to black-and-white, you will not be able to restore the colors.
3. Click the appropriate button for your task.
 - **Scan to Other Formats** scans a paper document and converts it into any of the supported formats
 - **File (PDF/Image) to Other Formats** converts a PDF document or image file into any of the supported formats
 - **Photo to Other Formats** converts a digital photo into any of the supported formats
4. A new document with the recognized text will be opened.

Important! When you run a quick task, data are converted based the program's current options. If you have changed application options, run the task again to recognize text with the new options.

Tip. If you do not see a task that suits your purposes, you can create your own automated task. See Automated Tasks for detailed instructions.

Custom Automated Tasks (only ABBYY FineReader Corporate Edition)

Custom automated tasks offer more processing steps than the built-in quick tasks. The steps in a task correspond to the document processing stages in ABBYY FineReader.

- How to Create a New Automated Task
- Task Steps
- Managing Automated Tasks


How to Create a New Automated Task

On the **My Tasks** tab of the **New Task** window, you can create your own tasks for your specific purposes.

1. Click the **Create...** button. If you already created tasks on this tab, the name of the button should change to **Create / Edit / View Tasks...**

Tip. You can also launch the **Task Manager** by selecting **Tools>Task Manager...**

2. On the **Task Manager** toolbar, click the **New** button and enter a name for your task.
3. In the left-hand pane of the **Task settings**, select one of the available steps. The selected step will be moved to the right-hand pane.
4. Select desired options for each step.

Note: Clicking the  button removes a step from the task. Some of the steps are dependent on other steps and may not be removed separately. For example, when you remove an **Analyze** step, the corresponding **Read** step will also be removed.

5. Once you have added all the steps and selected their options, click **Finish**.

Task Steps

To create your custom automated task, select one of the actions available for each step:

1. FineReader document

This is always the first step in an automated task. Once you select it, it will be moved to the right-hand pane.

Select one of the following commands:

- **Create new document**
A new FineReader will be created when the task is launched. Specify the options to be used when creating the document, processing images, and recognizing texts:
 - Use the document options selected in the program when the task is launched
 - Use custom document options

- **Select existing document**
Specify the FineReader document that contains the images to be processed.
- **Use open document**
Images from the opened FineReader document will be processed.

2. Acquire images

This step is optional if earlier you selected **Select existing document** or **Use open document**.

- **Open image / PDF**
Specify the folder from which ABBYY FineReader will take images when the task is launched.
- **Scan**
The scanning settings selected in ABBYY FineReader at the time of scanning will be used.

Note:

- a. If images are added to a document that already contains images, only the newly added images will be processed. If a new document is added, all images in the document will be processed.
- b. If an ABBYY FineReader document to be processed contains already recognized pages and already analyzed pages, the recognized pages will not be processed anew and the analyzed pages will be recognized.

3. Analyze

During this step, the program analyzes images to identify text areas, picture areas, table areas, or barcode areas. The **Analyze** step is added automatically when the **Read** step is selected.

- **Analyze the layout automatically, then adjust areas manually**
ABBYY FineReader will analyze the images and identify the appropriate areas on them.
- **Draw areas manually**
ABBYY FineReader will ask you to draw the appropriate areas manually.
- **Use an area template**
Specify the area template to be used by ABBYY FineReader to analyze the images. For more about area templates, see Using Area Templates.

4. Read

The options used for recognizing images depend on the options selected at the **FineReader document** step.

5. Save / Send

This step allows you to select the formats to which ABBYY FineReader should save the recognition results. You can also set up ABBYY FineReader to send the results to a certain application, send the acquired images or text by e-mail, or create a copy of the ABBYY FineReader document. An automated task may have more than one step of this type. Specify the saving options to be used:

- **Use the saving options selected in the program when the task starts**

- **Use custom saving options**

- **Save document**

Here you can specify the name of the file, its format, and the folder where the file should be stored.

Note: To avoid specifying a new folder each time the task is launched, select **Create a time-stamped subfolder**.

- **Send document**

Here you can select the application in which to open the resulting document.

- **E-mail document**

Here you can specify the name of the file, its format, and the e-mail address to which the file should be sent.

- **Save images**

Here you can specify the name of the file, its format, and the folder where the image file should be stored.

Note: To save all images to one file, select **Save as one multi-page image file** (only for TIFF and PDF).

- **E-mail images**

Here you can specify the name of the file, its format, and the e-mail address to which the image file should be sent.

- **Save FineReader document**

Here you can specify the folder to which the ABBYY FineReader document should be saved.

Managing Automated Tasks

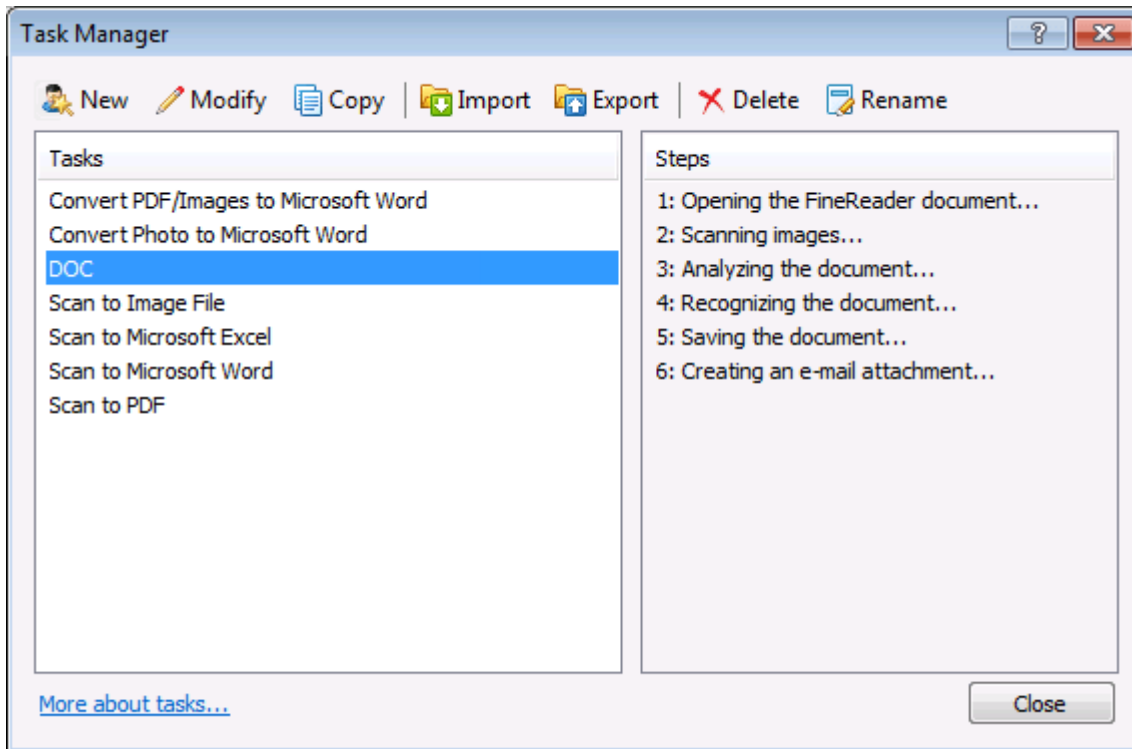
An ABBYY FineReader automated task is a set of successive steps. Each step corresponds to one of the document recognition stages. Similar actions can be automated using tasks.

Automated task management includes creating, editing, starting, and deleting custom automated tasks. Tasks are managed via the **Task Manager** dialog box.

To open the **Task Manager** dialog box, do one of the following:

- In the **New Task** window, click the **My Tasks** tab and then click **Create / Edit / View Tasks...**
- Select **Task Manager...** in the **Tools** menu
- Press CTRL+T

The toolbar at the top of the **Task Manager** contains buttons for creating, modifying, copying, deleting, importing, and exporting tasks.



To start processing a document, run the task on the **My Tasks** tab of the **New Task** window.

ABBYY FineReader Step-by-Step

Document processing in ABBYY FineReader has four stages: acquiring an image, recognizing it, checking the recognition results, and saving them. This section examines each of these steps in detail.

- Getting an Image
- Recognition
- Checking and Editing
- Saving Results
- E-mailing OCR Results
- Sending OCR Results to Kindle

Getting an Image

To begin the OCR process, ABBYY FineReader needs to retrieve an image of your document. There are several ways to create an image, including:

- Scanning your paper document
- Opening an existing image file or PDF document
- Taking a picture of your text

Scanning paper documents

1. Make sure that your scanner is connected and turned on.

Consult your scanner's documentation to ensure it is set up correctly. Be sure to install the software provided with your scanner. Some scanner models must be turned on before the computer is turned on.

2. Put the page to be scanned into the scanner. If the scanner is equipped with an automatic document feeder, you can put multiple pages into the scanner if you need to process a multi-page document.
3. Click **Scan**, or select **Scan Pages...** from the **File** menu.

The scanning dialog box will open. Select the appropriate settings and start scanning. The scanned pages will be added to the **Pages** window in ABBYY FineReader.

Tip. Typical printed office documents are best scanned in grayscale at 300 dpi.

The quality of OCR results depends on the quality of the source document and the selected scanning options. Low image quality may affect recognition results. To ensure optimal recognition accuracy, it is important to select the correct scanning options, taking into account the properties of your document.

Opening image files and PDF documents

You can open PDF documents and image files (for details, see Supported Image Formats).


How to open an image or PDF file:

- In **ABBYY FineReader**, click **Open**, or select **Open PDF File/Image...** from the **File** menu.

In the dialog box, select **Open image**:

1. Select one or more images.
2. Specify the page range for a multi-page document.
3. Enable the **Detect page orientation** option to automatically detect the page layout when opening images.
4. Activate the **Split facing pages** option to automatically split dual pages into separate images when opening images.
5. Enable the **Enable image preprocessing** option to perform automated image preprocessing when images are opened.

Note: When page images are opened or scanned for the first time, they are added to a new FineReader document, which is displayed in the **Pages** window. Any pages scanned or opened after this will be appended to the open FineReader document.

- In **Windows Explorer**, right-click the appropriate image file and select **Open in ABBYY FineReader** from the shortcut menu.
- In **Microsoft Outlook**, select the e-mail message with the image or PDF attachments you wish to open, and then click  on the toolbar. In the dialog box that opens, select the files to open.
- In **Microsoft Outlook** or **Windows Explorer**, drag the desired image file into ABBYY FineReader's main window. The image will be added to the current ABBYY FineReader document.

Note: The author of a PDF file may choose to restrict access to it. For example, the author may create a password or restrict certain features, such as the ability to extract text and graphics. To adhere to copyright guidelines, ABBYY FineReader will ask you for a password to open such files.

Taking a picture of a text with a camera

ABBYY FineReader can perform OCR on images created with a digital camera.

1. Take a picture of the text.

Note: Good OCR results can only be achieved with photos of good quality.

2. Save the photo to your hard disk.
3. Click the **Open** button or select **Open PDF File/Image...** from the **File** menu.

Recognition

ABBYY FineReader uses Optical Character Recognition technologies to convert document images into editable text. Prior to OCR, the program analyzes the structure of the entire document and detects the areas that contain text, barcodes, images, and tables.

By default, ABBYY FineReader documents are recognized automatically. The current program settings are used for automatic recognition.

Tip. You can disable automatic analysis and OCR for newly added images from the **Scan/Open** tab of the **Options** dialog box (**Tools>Options...**).

To achieve the best OCR quality, select the appropriate options for the recognition language, recognition mode, and recognized text print type.

In some cases, the OCR process can be started manually. For example, if you disabled automatic recognition, manually selected areas on an image, or changed the following settings in the **Options** dialog box (**Tools>Options...**):

- the recognition language on the **Document** tab
- the document type on the **Document** tab
- the color mode on the **Document** tab
- the recognition options on the **Read** tab
- the fonts to use on the **Read** tab

To launch the OCR process manually:

- Click the **Read** button on the main toolbar, or
- Select **Read** from the **Document** menu

Tip. To recognize the selected area or page, use the appropriate options on the **Page** and **Areas** menus, or use the shortcut menu.

Checking and Editing

Recognition results are displayed in the **Text** window. Unreliably recognized characters are highlighted in color in this window. This makes it easy to locate possible errors and quickly correct them.

You can edit the output document either directly in the **Text** window or with the built-in **Verification** dialog box (**Tools>Verification...**), which lets you review unreliably recognized words, find spelling errors, add new words into the dictionary, and change the dictionary language. ABBYY FineReader also lets you adjust the formatting of recognized text.

You can edit recognized text in the **Text** window by using the tools on the main toolbar and the **Text Properties** panel, which can be accessed by right-clicking inside the **Text** window and selecting **Properties** from the shortcut menu.

Note: When ABBYY FineReader performs OCR, it automatically detects the document's styles. All of the selected styles are displayed on the **Text Properties** panel. By editing styles, you can easily change the formatting of the entire text. ABBYY FineReader preserves all styles when saving text to RTF/DOC/DOCX/ODT formats.

Saving OCR Results

Recognition results can be saved to a file, sent to another application, copied to the Clipboard, or e-mailed. You can also send OCR results to Kindle.com where they will be converted to the Kindle format and become available for download onto your Kindle device. You can save either the entire document, or selected pages only.

Important! Be sure to select the appropriate saving options before clicking Save.

To save recognized text:

1. Select the format saving mode on the main toolbar's drop-down menu.

Document Format Saving Modes.

a. **Exact Copy**

Produces a document with formatting that corresponds to that of the original. Recommended for documents with complex formatting, such as promotion booklets. Note, however, that this option limits the ability to change the text and format of the output document.

b. **Editable Copy**

Produces a document in which the formatting may differ slightly from that of the original. Documents produced in this mode are easy to edit.

c. **Formatted Text**

Retains fonts, font sizes, and paragraphs, but does not retain the exact spacing or locations of the objects on the page. The text produced will be left-aligned. Texts in right-to-left scripts will be right-aligned.

Note: Vertical texts will be changed to horizontal in this mode.

d. **Plain Text**

This mode does not retain text formatting.

e. **Flexible layout**

Produces an HTML document with the positions of the objects as close to the original as is technically possible.

Important! The list of available options may vary depending on the selected saving format.

2. On the **Save** tab of the **Options** dialog box, select the desired saving options and click **OK**.
3. Click the arrow to the right of the **Save** button on the main toolbar and select the appropriate option, or use commands from the **File** menu.

Note: ABBYY FineReader lets you save page images. For details, see Saving an Image of the Page.

E-mailing OCR Results

With ABBYY FineReader, you can e-mail recognition results in any of the supported formats with the exception of HTML.

To send a document:

1. Select **E-mail>Document** from the **File** menu.
2. In the dialog box, select:
 - the format in which the document should be saved
 - one of the saving options: save all pages to one file, save each page to a separate file, or split the document into files based on blank pages
3. Click **OK**. An e-mail message will be created with the attached file(s). Simply specify the recipient and send the message.

You can also e-mail the images of document pages.

To send page images:

1. On the File menu, click E-mail>Page Images.
2. In the dialog box, select:
 - the format in which the document should be saved
 - **Send as multi-page image file** if you need to send images as a multi-page file
3. Click **OK**. An e-mail message will be created with the attached file(s). Simply specify the recipient and send the message.

Sending OCR Results to Kindle

With ABBYY FineReader, you can e-mail recognition results to the Kindle.com server, where they will be converted into the Kindle format and made available for download onto your Kindle device.

Important! To use this feature, you need a Kindle account on Kindle.com.

1. Click the arrow next to the **Save** button on the main toolbar and select **Send to Kindle...**
2. Specify:
 - your Account
 - one of the saving options: save all pages to one file, save each page to a separate file, or split the document into files based on blank pages
3. Click **OK**.
An e-mail message with an attachment will be created. Send this e-mail message to Kindle.com and the converted file will soon be available for download onto your Kindle device.

ABBYY FineReader Document

An ABBYY FineReader document contains both the images of the document pages and the recognized text. Most program settings (e.g. the scanning and OCR options selected by the user, patterns, languages, and language groups) apply to the entire ABBYY FineReader document.

- **General**
Provides information on how page images are displayed in the **Pages** window and how to customize the appearance of the main window.
- **Working with ABBYY FineReader Documents**
Provides information on how to open, delete, and close an ABBYY FineReader document, how to add images to a document, and how to save a document or document options.
- **Splitting a FineReader document**
- **Numbering Pages in an ABBYY FineReader Document**

General

A new ABBYY FineReader document is opened by default when you start ABBYY FineReader. You can either start working with this new document or open an existing ABBYY FineReader document.

All document pages are displayed in the **Pages** window. The pages in the **Pages** window can be displayed either as thumbnails or as page icons.

To change the display mode:


- Click View>Pages Window. Then select Thumbnails or Details.
- On the shortcut menu of the **Pages** window, select **Pages Window** and then select either **Thumbnails** or **Details**.
- On the **View** tab of the **Options** dialog box (**Tools>Options...**), in the **Pages Window** group, select either **Thumbnails** or **Details**.

To open a page, click its thumbnail in the **Pages** window (in **Thumbnails** mode) or double-click its number (in **Details** mode). In the **Image** window you will see the page image and in the **Text** window you will see the recognition results (if the document was recognized).

Note: When opening FineReader documents created in earlier versions of ABBYY FineReader, the program attempts to convert them into its new format. As this conversion cannot be undone, ABBYY FineReader will prompt you to save the converted document under a new name. All recognized text will be lost in the process.

Working with ABBYY FineReader Documents

When working with an ABBYY FineReader document, you can:

- Create a new document
 - On the File menu, click New FineReader Document, or
 - On the main toolbar, click .
- Delete a page from the current document
 - In the **Pages** window, select the page you wish to delete and then select **Delete Page from Document** from the **Page** menu, or

- In the **Pages** window, right-click the page you wish to delete and then select **Delete Page from Document** from the shortcut menu, or
- In the **Pages** window, select the page you wish to delete and press the DELETE key.

To select several pages, hold down the CTRL key and click each of the pages to delete one by one.


- Open a document

At launch, ABBYY FineReader automatically creates a new document.

Note: If you wish ABBYY FineReader to open the last document on which you were working in the previous session, select **Tools>Options...**, click the **Advanced** tab, and then select **Open the last used FineReader document when the program starts**.

To open a document:

1. On the File menu, click Open FineReader Document...
2. In the **Open Document** dialog box, select the document to open.

Note: You can also open an ABBYY FineReader document directly from Windows Explorer by right-clicking the desired document folder (document folders are displayed as ) and selecting **Open with ABBYY FineReader** from the shortcut menu.

- Add images to the current document

- On the File menu, click Open PDF File/Image...
- In the **Open Image** dialog box, select the desired images and click **Open**. The image will be appended to the end of the current document and its copy will be saved in the folder that stores the document.

Note: You can also add images directly from Windows Explorer:

- In Windows Explorer, select the desired image files, right-click the selection, and select **Open with ABBYY FineReader** from the shortcut menu.

If ABBYY FineReader is running, the selected files will be added to the current document. Otherwise, ABBYY FineReader will be launched and the selected files will be added to a new document.

- Save the current document

1. On the File menu, click Save FineReader Document...
2. In the **Save Document** dialog box, type in a name for your document and specify a storage location.

Note: Saving an ABBYY FineReader document will save the page images, the recognized text, the trained patterns, and the user languages (if any).

- Close the current document

- To close a page of the current document, select the page and then select **Close Current Page** from the **Document** menu.
- To close the entire document, select **Close FineReader Document** from the **File** menu.

- Save user patterns and languages

To save your user patterns and languages in one file:

1. Select **Tools>Options...** and click the **Read** tab.
2. Under User patterns and languages, click the Save to File... button.
3. In the **Save Options** dialog box, type in a name for your file and specify a storage location.

This file will contain the path to the folder where user languages, language groups, dictionaries, and patterns are stored.

- Load user patterns and languages

1. Select **Tools>Options...** and click the **Read** tab.
2. Under User patterns and languages, click the Load from File... button.
3. In the **Load Options** dialog box, select the file that contains the desired user patterns and languages and click **Open**.

Splitting a FineReader Document

When processing large numbers of multi-page documents, it is often more practical to scan all the documents first and only then analyze and recognize them. However, to preserve the original formatting of each paper document correctly, ABBYY FineReader must process each of them as a separate FineReader document. ABBYY FineReader offers tools to group scanned pages into separate documents to which they belong.

To split an ABBYY FineReader document into several documents:

1. On the **File** menu, click **Split into Several Documents...** or select pages in the **Pages** window, right-click the selection, and then select **Create New Document with Selected Pages...**
2. In the **Split Document** dialog box, click **Add document** to create the required number of FineReader documents.
3. Move pages from the **Pages** pane into their appropriate documents displayed in the **New Documents** pane using one of the following three methods:
 - Select pages and drag them with the mouse.
Note: You can also use drag-and-drop to move pages between documents.
 - Click the **Move** button to move the selected pages into the current document displayed in the **New Documents** pane or click the **Return** button to remove the selected pages from the current document.
 - Press the **CTRL+RIGHT ARROW** to move pages into the current document and **CTRL+LEFT ARROW** or **DELETE** to remove the selected pages from the current document.
4. Once you are finished moving pages into the new FineReader documents, click the **Create All** button to create all documents at once or click the **Create** button in each of the documents individually.

Note: You can also drag-and-drop selected pages from the **Pages** window into any other ABBYY FineReader window. A new ABBYY FineReader document will be created for these pages.

Numbering Pages in an ABBYY FineReader Document

All pages in a document are numbered. Page numbers are displayed in the **Pages** window next to the page thumbnails.

To renumber pages:

1. In the **Pages** window, select one or multiple pages.
2. On the shortcut menu of the **Pages** window, click **Renumber Pages....**
3. In the **Renumber Pages** dialog box, choose one of the following:
 - **Renumber pages (cannot be undone)**
This changes all page numbers successively, starting with the selected page.
 - **Restore original page numbers after duplex scanning**
This option restores the original page numbering of a document with double-sided pages if you used a scanner with an automatic feeder to first scan all the odd-numbered pages and then all the even-numbered pages. You can choose between the normal and the reverse order for the even-numbered pages.

Important! This option will only work for 3 or more consecutively numbered pages.
 - **Swap book pages**
This option is useful if you scan a book written in a left-to-right script and split the facing pages, but fail to specify the correct language.

Important! This option will only work for 2 or more consecutively numbered pages, including at least 2 facing pages.
4. Click **OK**.

The order of the pages in the **Pages** window will change to reflect the new numbering.

Note:

1. To change the number of one page, click its number in the **Pages** window and enter the new number in the field.
2. In **Thumbnails** mode, you can change page numbering simply by dragging selected pages to the desired place in the document.

Taking Into Account Some of the Features of Your Paper Document

OCR quality greatly depends on the quality of the source image. Consider the following elements to ascertain whether you will get the scanning results you desire:

- **Document Type**
Various devices may be used to produce printed documents, such as typewriters or faxes. Recognition quality for these documents may vary. This section provides recommendations for selecting the right print type.
- **Print Quality**
OCR quality may be greatly impaired by "noise" that sometimes occurs on poor quality documents. This section provides recommendations for scanning these documents.

- **Color mode**
A document may be color or black-and-white. More about selecting a color mode for OCR.
- **Document Languages**
A document may contain text written in multiple languages. This section provides recommendations for selecting recognition languages.

Document Type

When recognizing typewritten texts or faxes, OCR quality can be improved by selecting the corresponding **Document type** in the **Options** dialog box.

For most documents, the program will detect the print type automatically. For automatic print type detection, the **Auto** option must be selected under **Document type** in the **Options** dialog box (the **Tools>Options...>** menu, then the **Document** tab). You can process the document in full-color or black-and-white mode.

You may also choose to manually select the print type as needed.

	<p>An example of typewritten text. All letters are of equal width (compare, for example, "w" and "t"). For texts of this type, select Typewriter.</p>
	<p>An example of a text produced by a fax machine. As you can see from the example, the letters are not clear in some places, in addition to noise and distortion. For texts of this type, select Fax.</p>

Tip. After recognizing typewritten texts or faxes, be sure to select **Auto** before processing regular printed documents.

Print Quality

Poor-quality documents with "noise" (i.e. random black dots or speckles), blurred and uneven letters, or skewed lines and shifted table borders may require specific scanning settings. Show me an example...

Fax

Newspaper

Poor-quality documents are best scanned in grayscale. When scanning in grayscale, the program will select the optimal brightness value automatically.

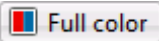
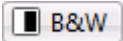
Grayscale scanning mode retains more information about the letters in the scanned text to achieve better OCR results when recognizing documents of medium to poor quality. You can also correct some of the defects manually using the image editing tools available in the Image Editor. For details, see *Editing Images Manually*.

Color Mode

If you do not need to preserve the original colors of a full-color document, you can process the document in black-and-white mode. This will greatly reduce the size of the resulting ABBYY FineReader document and speed up the OCR process. However, processing low-contrast images in black and white may result in poor OCR quality. We also do not recommend black and white processing for photos, magazine pages, and texts in Chinese, Japanese, and Korean.

Note: You can also speed up recognition of color and black-and-white documents by selecting the **Fast reading** option on the **Read** tab of the **Options** dialog box. For more about the recognition modes, see *OCR Options*.

To select a color mode:

- Use the **Color mode** drop-down list in the **New Task** window or
- Use the options under **Color mode** on the **Document** tab of the **Options** dialog box (**Tools>Options...**) or
- Click  or  in the **Pages** window.

Important! Once the document is converted to black-and-white, you will not be able to restore the colors. To get a color document, open the file with color images or scan the paper document in color mode.

Document Languages

ABBYY FineReader recognizes both single- and multi-language documents (e.g. written in two or more languages). For multi-language documents, you need to select several recognition languages.

To specify an OCR language for your document, in the **Document Languages** drop-down list on the main toolbar or in the **New Task** window, select one of the following:

- **Autoselect**
ABBYY FineReader will automatically select the appropriate languages from the user-defined list of languages. To modify this list:
 1. Select **More languages...**
 2. In the Language Editor dialog box, select the option **Automatically select document languages from the following list**.
 3. Click the **Specify...** button.
 4. In the **Languages** dialog box, select the desired languages.
- **A language or a combination of languages**
Select a language or a language combination. The list of languages includes the languages most often used on the computer, as well as English, German, and French.
- **More languages...**
Select this option if the language you need is not visible in the list.

In the **Language Editor** dialog box, select the **Specify languages manually** option and then select the desired language or languages by checking the appropriate boxes. If you often use a particular language combination, you can create a new group for these languages.

There are several reasons that a language may not be listed:

1. The language is not supported by ABBYY FineReader.
2. Your copy of the software product does not support this language.

The complete list of languages available in your copy can be found in the **Licenses** dialog box (select **Help > About... > License Info**).

Image Acquisition Tips

OCR quality depends on the quality of the original image and the settings used during the scanning process. This section provides recommendations on scanning documents and taking photos of documents with a digital camera. There is also some advice on correcting defects that commonly occur on scanned images and digital photos.

- Customizing Scanning and Image Opening
- Selecting the Scanning Interface
- Scanning Recommendations
- Taking Photos of Documents
- Automatic Image Preprocessing
- Editing Images Manually

Customizing Scanning and Image Opening

In ABBYY FineReader, you can select both global processing options and options for scanning and opening individual pages. For example, you can select the scanning interface and enable/disable automatic document analysis, OCR, and image preprocessing.

You can select the required options either directly from the image opening or scanning dialog boxes (if you are using the ABBYY FineReader interface to scan documents) or on the **Scan/Open** tab of the **Options** dialog box (**Tools>Options...**).

Important! After you change any settings in the **Options** dialog box, you will need to re-scan or re-open your images.

The following settings are available on the **Scan/Open** tab of the **Options** dialog box:

- Launching automatic analysis and recognition of the retrieved images

ABBYY FineReader is by default set to analyze and recognize documents automatically. These settings can be modified as needed. The following options are available:

- **Automatically read acquired page images**
Document analysis and recognition will be performed automatically.
- **Automatically analyze acquired page images**
Document analysis will be performed automatically, and recognition will require the user's attention.

- **Do not read and analyze acquired page images automatically**
The scanned or opened images will be added to the FineReader document. The user will have to initiate document analysis and recognition manually. This mode is used for processing documents with a complex structure.
- Image processing methods
 - **Enable image preprocessing**
Enable image preprocessing if you need to OCR a book or a digital photo of a text document. When activated, this feature automatically decides what action to take depending on the input image type and applies the correction. Adjustments include removing noise from digital photos, deskewing, removing motion blur, straightening text lines, and correcting trapezium distortions.
 - **Detect page orientation**
Enable this feature to automatically detect the orientation of pages added to the FineReader document.
 - **Split facing pages**
Enable this feature when scanning a facing or opening dual pages. It will automatically split pages before adding them to the FineReader document.

Note: The image preprocessing features are only available in the Image Editor dialog box (**Page>Edit Page Image...**) to process the images opened in FineReader, and are not available while scanning or opening document pages. For details, see "Editing Images".

- The scanner interface to be used while scanning

Selecting the Scanning Interface

By default, ABBYY FineReader's scanning dialog box is used to adjust the scanning options. If, for some reason, the ABBYY FineReader scanning dialog box is incompatible with your scanner model, the scanner's dialog box is used.

1. ABBYY FineReader scanning dialog box

In this case, the **ABBYY FineReader** interface is used to set up scanning mode. The following settings are available:

- Resolution, Brightness, and Scanning Mode
- Image preprocessing options:
 - a. **Detect page orientation** (enable this feature to automatically detect page orientation while scanning)
 - b. **Split facing pages** (enable this feature to automatically split facing pages into separate images)
 - c. **Enable image preprocessing** (enable this feature to perform automatic processing of scanned images)
- Multi-page document scanning options:
 - a. Use the Automatic Document Feeder (ADF)
 - b. Apply automatic duplex scanning (if supported by the scanner)
 - c. Set the page scanning delay in seconds

Note: The **Use ABBYY FineReader interface** option may not be available for some scanner models.

2. TWAIN or WIA interface provided by the scanner

In this case, scanning options are selected from the scanner driver's dialog box. Consult the technical documentation provided with the scanner for more information about the dialog box and its components.

Important! Consult the scanner manual to make sure that you have properly set up your scanner. Make sure to install the software supplied with the scanner.

The ABBYY FineReader scanning interface is used by default.

To change the scanning interface:

1. Open the **Options** dialog box from the **Scan/Open** tab (**Tools>Options...**).
2. Set the toggle switch in the **Scanner** group to either **Use ABBYY FineReader interface** or to **Use native interface**.

Scanning Tips

You can set the scanning options in the ABBYY FineReader scanning dialog box or in the scanner driver dialog box. See the "Selecting the Scanning Interface" section for more information about the dialog box.

- Make sure that your scanner is connected properly and switched on.
Consult your scanner documentation to ensure it is set up correctly. Be sure to install the software provided with your scanner. Some scanner models must be switched on before you power up the computer.
- Place your documents face down on the scanner, making sure they are properly adjusted.
Skewed images may be converted incorrectly.
- Printed documents are best scanned in grayscale at 300 dpi.
Conversion quality depends on the quality of the original document and scanning parameters. Poor image quality may have an adverse effect on the quality of conversion. Be sure to select the scanning parameters appropriate for your document.

More information about some of the scanning options is available in these sections:

- Tuning Brightness
- Font is Too Small
- Adjusting Image Resolution
- Scanning Facing Pages

Tuning Brightness

ABBYY FineReader will display a warning message during scanning if the brightness setting is incorrect. You may also need to adjust the brightness setting when scanning in black-and-white mode.

To adjust the brightness:

1. Click the **Scan** button.

2. In the dialog box, specify the desired brightness.
Depending on which scanning interface is being used, either the ABBYY FineReader scanning dialog box or the scanner driver dialog box will open. A medium value of around 50% should suffice in most cases.
3. Proceed to scan the document.

If the resulting image contains too many "torn" or "glued" letters, troubleshoot using the table below.

Your image looks like this	Recommendations
brightness	This image is suitable for OCR.
brightness characters are "torn" or very light	<ul style="list-style-type: none"> Lower the brightness to make the image darker. Scan in grayscale. Brightness will be tuned automatically.
brightness characters are distorted, glued together, or filled	<ul style="list-style-type: none"> Increase the brightness to make the image brighter. Scan in grayscale. Brightness will be tuned automatically.


Font is Too Small

For optimal OCR results, scan documents printed in very small fonts at higher resolutions.

1. Click the **Scan** button.
2. In the dialog box, specify the desired resolution.
Depending on which scanning interface is being used, either the ABBYY FineReader scanning dialog box or the scanner driver dialog box will open.
3. Proceed to scan the document.

You may wish to compare the images of the same document obtained at different resolutions by opening them in the **Zoom** window in **Pixel-to-Pixel** mode (**View>Zoom Window>Scale>Pixel-to-Pixel**).

Source image	Recommended resolution
FineReader	300 dpi for typical texts (printed in fonts of size 10pt or larger)

	400–600 dpi for texts printed in smaller fonts (9pt or smaller)
---	---

Adjusting Image Resolution

OCR quality is affected by the resolution at which documents are scanned. Low resolutions may adversely affect the quality of OCR.

We recommend scanning documents at 300 dpi.

Important! For best OCR results, vertical and horizontal resolutions must be the same.

Setting the resolution too high (over 600 dpi) increases the recognition time. Increasing the resolution does not yield substantially improved recognition results. Setting an extremely low resolution (less than 150 dpi) adversely affects OCR quality.

You may need to adjust the resolution of your images if:

- The resolution of your image is less than 250 dpi or higher than 600 dpi.
- The resolution of your image is non-standard.
For example, faxes may have a resolution of 204 x 96 dpi.

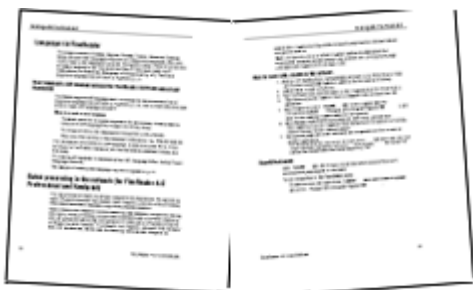
To adjust the resolution:

1. Click the **Scan** button.
2. Depending on the scanning interface used, the ABBYY FineReader interface or the scanner driver interface will appear.
Choose the desired resolution in the scanning dialog box.
3. Scan the document.

Tip. You can also adjust the resolution in the **Image Editor (Page>Edit Page Image...)**.

Scanning Facing Pages

When scanning facing pages from a book, both pages will appear in a single image. See Sample Image.



To improve OCR quality, split the facing pages into two separate images. The program features a special mode that automatically splits two facing pages into two separate images in the ABBYY FineReader document.

Scanning facing pages or dual pages:

1. Open the **Options** dialog box from the **Scan/Open** tab (**Tools>Options...**).

2. Check the Split facing pages option in the Image Processing group.
3. Scan the pages.

The automatic splitting option is also available in the **image opening dialog (File>Open PDF File/Image...)** and ABBYY FineReader's **scanning dialog box**. For details, see Selecting the Scanning Interface.

Important! If the image contains text in Chinese or Japanese, use the **Edit Image** dialog box to split pages manually. These types of images are not supported by the automatic page splitting feature.

To split facing pages manually:

1. Open the Edit Image dialog box (Page>Edit Page Image...).
2. Use the options in the **Split** group to split the page.

Taking Photos of Documents

- Camera Requirements
- Lighting
- Taking Photos
- When you need to take another photo

Camera Requirements

Your camera should meet the following requirements in order to obtain document images that can be reliably recognized.

Recommended camera characteristics

- Image sensor: 5 million pixels for A4 pages. Smaller sensors may be sufficient for taking pictures of smaller documents such as business cards.
- Flash disable feature
- Manual aperture control, i.e. availability of Av or full manual mode
- Manual focusing
- An anti-shake system or ability to use a tripod
- Optical zoom

Minimum requirements

- Image sensor: 2 million pixels for A4 pages
- Variable focal distance. Fixed-focus cameras (lens focused on infinity) should not be used to photograph documents. These types of camera are commonly built into mobile phones and PDAs.

Note: For detailed information about your camera, please refer to the documentation supplied with your device.

Lighting

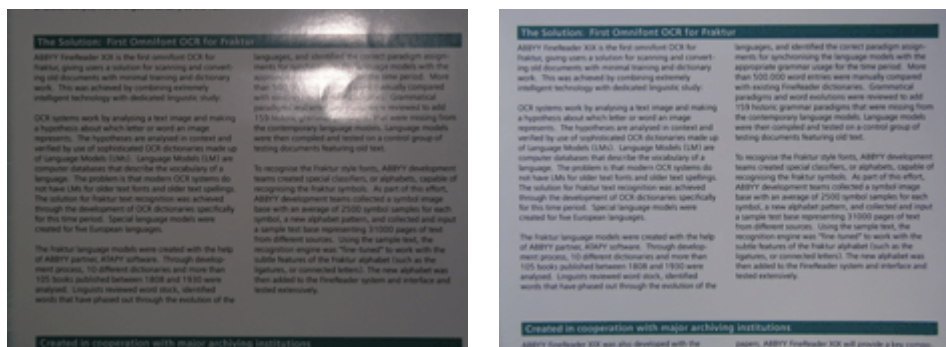
Lighting greatly affects the quality of the resulting photo.

Best results can be achieved with bright and evenly distributed light, preferably daylight. On a bright sunny day, you can increase the aperture number to get a sharper picture.

Using flash and additional light sources

- When using artificial lighting, use two light sources positioned so as to avoid shadows or glare.
- If there is enough light, turn the flash off to prevent sharp highlights and shadows. When using the flash in poor lighting conditions, be sure to take photos from a distance of approximately 50 cm.

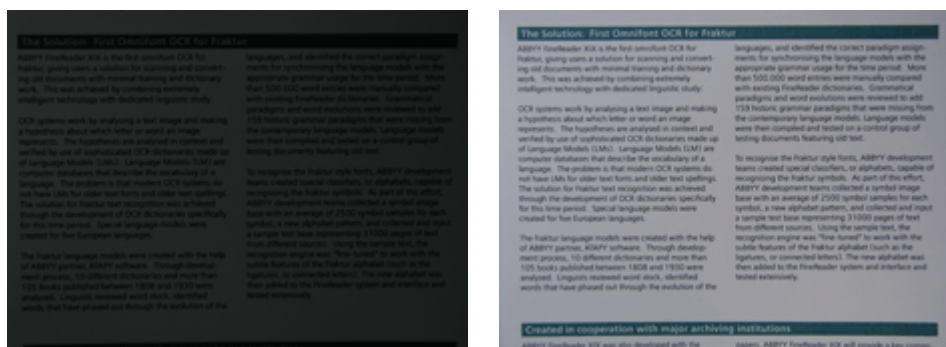
Important! The flash must not be used to take pictures of documents printed on glossy paper. Compare an image with glare with a good quality image:



If the image is too dark

- Set a lower aperture value to open up the aperture.
- Set a higher ISO value.
- Use manual focus, as automatic focus may fail and cause blur.

Compare an image that is too dark with a good quality image:



Taking Photos

To obtain good quality photos of documents, be sure to position the camera correctly and follow these simple recommendations.

- Use a tripod whenever possible.

- The lens should be positioned parallel to the page. The distance between the camera and the document should be selected so that the entire page fits within the frame when you zoom in. In most cases this distance will be between 50 and 60 cm.
- Even out the paper document or book pages (especially in the case of thick books). The text lines should not be skewed by more than 20 degrees, otherwise the text may not be converted properly.
- To get sharper images, focus on the center of the image.



- Enable the anti-shake system, as longer exposures in poor lighting conditions may cause blur.
- Use the automatic shutter release feature. This will prevent the camera from moving when you press the shutter release button. The use of automatic shutter release is recommended even if you use a tripod.

When you need to take another photo

Please make sure that your camera meets the technical requirements and follow the simple recommendations provided in the "Lighting" and "Taking Photos" sections.

How to improve an image, if:

- the image is too dark or its contrast is too low
Solution: Try to use brighter lighting. Otherwise, set a lower aperture value, i.e. open up the aperture.
- the image is not sharp enough
Solution: Autofocus may not work properly in poor lighting or when taking pictures from a close distance. Try using brighter lighting. If this does not help, try focusing the camera manually.
 If the image is only slightly blurred, try the **Photo Correction** tool that is available in the Image Editor. For more information, see the "Editing Images Manually" section.
- only part of the image is blurry
Solution: Try setting a higher aperture value. Take pictures from a greater distance at maximum optical zoom. Focus on a point between the center and the edge of the image.
- flash causes glare
Solution: Turn off the flash or try using other light sources and increasing the distance between the camera and the document.

Automatic Image Preprocessing

OCR quality may be affected by distorted text lines, document skew, noise, and other defects commonly found on scanned images and digital photos.

ABBYY FineReader offers an image preprocessing feature to deal with image defects of this kind. When activated, the image preprocessing feature automatically decides what action to take depending on the input image type and applies the correction. Adjustments include removing noise from digital photos, deskewing, straightening text lines, and correcting trapezium distortions.

Note: These operations can be very time consuming.

To activate automatic preprocessing when images are scanned and opened:

1. Open the **Options** dialog box (**Tools>Options...**).
2. Make sure to check the **Enable image preprocessing** option on the **Scan/Open** tab in the **Image Processing** group.

Important! The automatic preprocessing option may also be toggled in the image opening dialog box (**File menu>Open PDF File/Image...**) and in the ABBYY FineReader scanning dialog box. For more information, see Selecting the Scanning Interface.

Note: If the **Enable image preprocessing** option was turned off while scanning or opening images, the image can be processed manually using the Edit Image dialog box. For details, see Editing Images Manually.

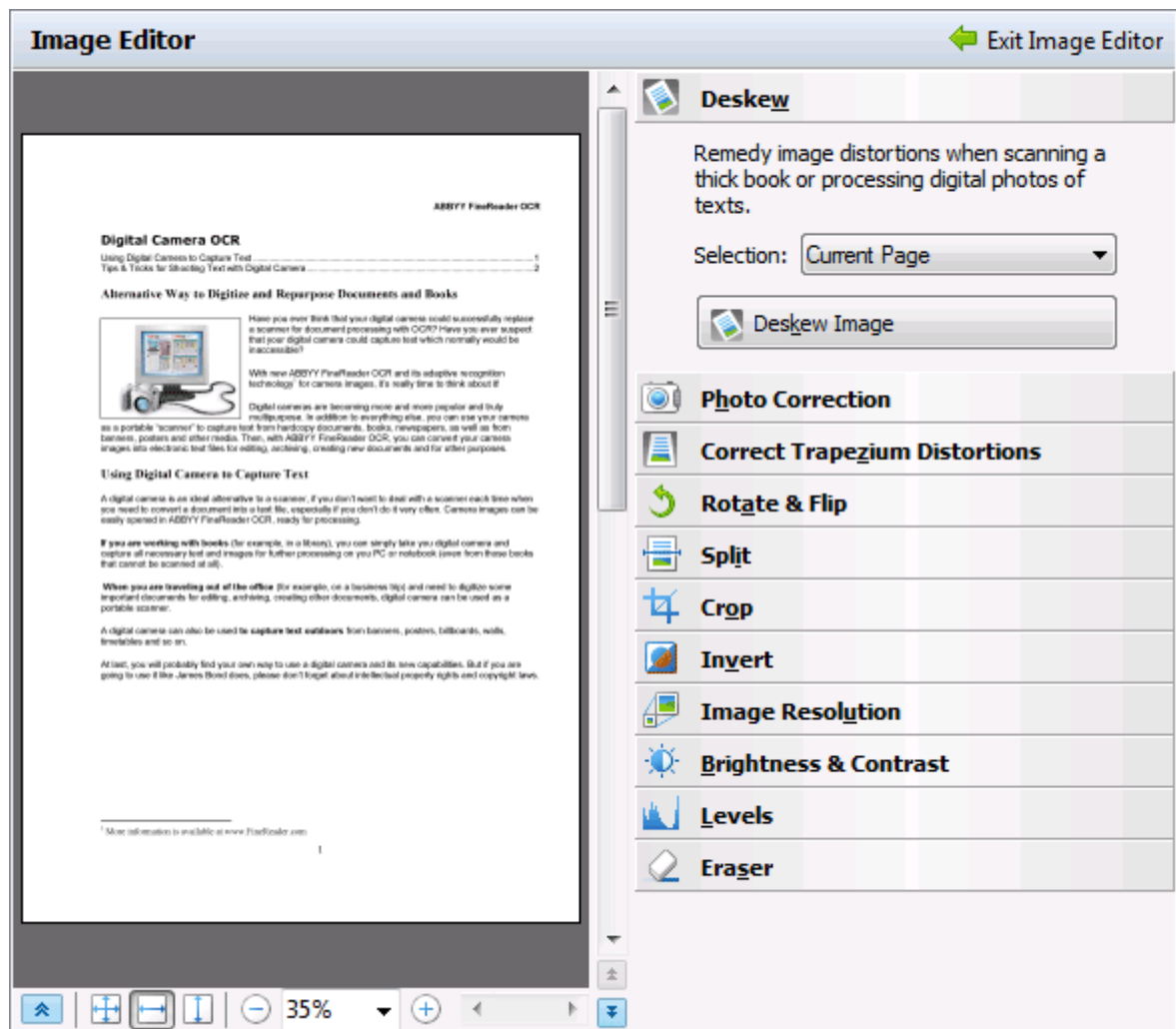
Editing Images Manually

ABBYY FineReader 11 offers automatic image preprocessing, which means that the program decides if any correction is needed and applies it to the image.

However, some images may require manual editing. For example, when only a part of an image needs to be recognized or if you need to process only some of the pages in a document.

To edit an image:

1. Select **Page>Edit Page Image...**



In the **Image Editor**, the active page of the current ABBYY FineReader document is displayed on the left and the image editing tools are shown on the right.

2. An overview of the image editing tools:

- **Deskew** corrects image skew.
- **Photo Correction** straightens text lines, removes motion blur, and reduces noise.
- **Correct Trapezium Distortions** corrects perspective distortions. When you select this tool, a frame with grid lines is superimposed on the image. Drag and drop the corners of the frame to position them in the corners of the page image. Make sure that the horizontal grid lines run parallel to the lines of text. Then click **Correct This Image**.
- **Rotate & Flip** rotates the image and flips it left to right or right to left so that the text appears in the normal reading position (i.e. lines are horizontal and are printed left to right).
- **Split** splits the image (e.g. facing pages) into separate units.
- **Crop** is used to crop the unneeded edges of an image.

- **Invert** enables standard inversion of document colors (dark text against light background).
- **Image Resolution** changes image resolution.
- **Brightness & Contrast** allows you to select the optimal brightness and contrast values to improve image quality.
- **Levels** allows you to adjust the tonal values of the image by selecting the levels for shadows, highlights, and midtones on a histogram.
To increase image contrast, move the right and left sliders on the input levels histogram. The tone corresponding to the position of the left slider will be assumed to be the blackest part of the image, and the tone corresponding to the position of the right slider will be assumed to be the whitest part of the image. The remaining levels between the sliders will be distributed between level 0 and level 255. Moving the central slider to the right or to the left will make the image darker or brighter respectively.
To decrease image contrast, adjust the sliders for the output levels.
- **Eraser** erases a part of the image.

Click the button of the tool you want to use. Most of the tools can be applied either to all of the pages or only to the selected pages. The pages to edit can be specified using the **Selection** drop-down list of the tool or selected in the **Pages** window.

3. Once you are finished editing the image, click **Exit Image Editor**.

Tips for Improving OCR Quality

This section provides advice on selecting the OCR options that best match the structure of your document and on choosing the right settings to handle non-standard situations.

- OCR Options
- When the Complex Structure of a Paper Document Is Not Reproduced in the Electronic Document
- Adjusting Area Shapes and Area Borders
- Picture Not Detected
- Barcode Not Detected
- Table Not Detected
- Table Cells Detected Incorrectly
- Adjusting Text Area Properties
- Vertical or Inverted Text Not Recognized Properly
- Paper Document Contains Decorative (Non-Standard) Fonts
- Incorrect Font in Recognized Text or Some Characters Are Replaced with "?" or "□"

OCR Options

Selecting the right OCR options is important if you want fast, accurate results. For flawless and fast OCR. When deciding which options you want to use, you should consider not only the type and complexity of your document, but also how you intend to use the results.

You can find the OCR options on the **Read** tab of the **Options** dialog box (**Tools>Options...**).

Important! ABBYY FineReader automatically recognizes any pages you add to a FineReader document. The currently selected options will be used for recognition. You can turn off automatic analysis and OCR of newly added images on the **Scan/Open** tab of the **Options** dialog box (**Tools>Options...**).

Note: If you change the OCR options after a document has been recognized, run the OCR process again to recognize the document with the new options.

You can change the following OCR options on the **Read** tab of the **Options** dialog box:

- **Reading mode**

Select a reading mode.

Two reading modes are available:

- **Thorough reading**

This is a default mode. In this mode, ABBYY FineReader analyzes and recognizes both simple documents and documents with complex layouts, even those with text printed on a color or raster background and documents with complex tables (including tables with white grid lines and tables with color cells).

Note: Compared to the **Fast** mode, the **Thorough** mode takes more time but ensures better recognition quality.

- **Fast reading**

This mode is recommended for processing large documents with simple layouts and good quality images.

Select the mode that best suits your needs.

- **Training**

Pattern training is disabled by default. Select the **Read with training** option to teach ABBYY FineReader new characters when it performs OCR on your pages.

Pattern training is useful for the following types of texts:

- Texts printed in decorative fonts
- Texts which contain special characters (e.g. mathematical symbols)
- Large documents (more than 100 pages) of poor print quality

You can either use the built-in character patterns or create your own. Select the desired option under **Training**.

- **User patterns and languages**

You can save user patterns and languages to a file and load previously saved user patterns and languages.

The user pattern and language file you create stores the path to the folder with your user languages, language groups, dictionaries, and patterns.

To save user patterns and languages:

1. Click the **Save to File...** button
2. In the **Save Options** dialog box, enter a name for the file and click **Save**.

To load user patterns and languages:

1. Click the **Load from File...** button.
2. In the **Load Options** dialog box, select the desired *.fbt file and click **Open**.

- **Fonts**

Here you can select the fonts to be used when saving recognized text.

To select fonts:

1. Click the **Fonts** button.
2. Select the desired fonts and click **OK**.

- **Other**

If your document contains barcodes and you wish them to be converted into strings of letters and digits rather than saved as pictures, select **Look for barcodes**. This option is cleared by default.

If the complex structure of a paper document is not reproduced in the electronic document

Before running the document recognition process, ABBYY FineReader conducts an analysis of the logical structure of the document and detects areas containing text, pictures, tables, and barcodes. The program then relies on this analysis to identify the areas and the recognition order. This information helps recreate the formatting of the original document.

By default, ABBYY FineReader will automatically analyze documents automatically. However, in the case of very complex formatting some areas may be detected incorrectly and you may wish to correct them manually.

Use the following tools and commands to correct areas:

- Manual area adjustment tools on the **Image** window toolbar.
- Shortcut menu commands available by right-clicking the selected area in the **Image** window or the **Zoom** window.
- Commands in the main **Areas** menu.

Restart the OCR process after making the desired adjustments.

- Adjusting Area Shapes and Area Borders — area tools.
- Table Not Detected — the **Table** area and guidelines for turning the table on a page into a separate area.
- Table Cells Detected Incorrectly — tools for manual table markup.
- Picture Not Detected — the **Picture** area and guidelines for separating the image on a page into a dedicated area.
- Barcode Not Detected — the **Barcode** area and guidelines for separating the barcode on a page into a dedicated area.

Adjusting Area Shapes and Area Borders

ABBYY FineReader analyses page images before recognition and detects different types of areas on these images, such as **Text**, **Image**, **Table**, and **Barcode** areas. This analysis determines which areas will be recognized and in what order. In the user interface, different area types are distinguished by the colors of their borders.

If you click on an area, it will be highlighted and will become active. You can also navigate areas by pressing the TAB key. The areas are numbered so that you can see which area will be selected next when you press TAB. By default, these area numbers are not visible in the Image window, and only appear if the **Order Areas** tool is selected.

If the text of a non-active area does not fit entirely within its borders (this may happen, for example, after you edit the recognized text), some of the text may not be visible on the screen. Red marks that appear on the borders of the area will alert you to this. Once an area becomes active, its borders will expand to enclose all of the text.


The tools for drawing and editing areas manually can be found in the **Image** window and on the pop-up toolbars for the **Text**, **Image**, **Background Picture**, and **Table** areas (pop-up toolbars appear next to the active area).

Important! After you have finished editing the areas, run the OCR process again.

Creating a new area


1. Select a tool from the **Image** window:

 draws a Recognition area

 draws a Text area

 draws a Picture area

 draws a Background Picture area

 draws a Table area

2. Use the mouse to draw an area.

New areas can be drawn without selecting a tool from the **Image** window toolbar. You can simply draw an area on the image using one of the following key combinations:

CTRL+SHIFT draws a **Text** area

ALT+SHIFT draws an **Image** area

CTRL+ALT draws a **Table** area

CTRL+SHIFT+ALT draws a **Barcode** area



You can change the area type. Right-click the selected area to choose **Change Area Type** from the shortcut menu, and then select the desired area type.

Adjusting area borders

1. Place the mouse cursor on the area border.
2. Left-click and drag in the desired direction.
3. Release the mouse button when done.

Note: You can simultaneously adjust the vertical and horizontal borders of the area by placing the mouse pointer on a corner of the area.





Adding/removing area parts

1. Select the  /  tool on the pop-up toolbar for **Text**, **Image** or **Background Picture** areas.
2. Place the mouse cursor inside the area and select part of the image. This segment will be added to or deleted from the area.
3. Adjust the area border, if necessary.

Notes:

1. These tools may only be applied to **Text**, **Image** or **Background Picture** areas. You cannot add/delete a segment in a table or barcode area.
2. The area borders can also be adjusted by adding vertices (splitting points). The resulting segments can be dragged in any direction using the mouse. To add a new vertex, move the cursor to the desired point on the border while holding down CTRL+SHIFT (this switches the pointer to crosshairs) and click. A new vertex will appear on the area border.

Selecting one or more areas

- Select the , ,  or  tool on the **Image** window toolbar and click the desired areas while holding down the CTRL key. To cancel the selection, click the selected area once again.

Moving an area


- Select an area and move it while holding down the CTRL key.

Re-numbering areas

1. Select **Order Areas** from the **Areas** menu.
2. Select the areas in the order desired for the output document.

Note: The areas are automatically numbered from left to right during automatic page analysis, regardless of the direction of the text on the image.

Deleting an area

- Select the  tool and click the area you wish to delete.
- Select the areas to delete and then click **Delete Area** from the shortcut menu.
- Select the areas to delete and then press the **DELETE** key.

To delete all areas:

- Select **Delete All Areas and Text** from the **Image** window shortcut menu.

Important! Deleting an area from an image that has already been recognized will delete all of the text in the corresponding **Text** window.

Changing text orientation

- Select the desired text orientation from the drop-down **Orientation** menu on the **Area Properties** toolbar (the **Area Properties** shortcut menu).


For details, see Changing Text Properties.

Changing area border colors


1. Open the **Options** dialog box in the **View** tab (**Tools>Options...**).
2. Select the area type under **Appearance** and choose the color on the **Color** palette.

Picture Not Detected

Picture areas mark the pictures contained in your document. You can also use Picture areas if you do not want a text fragment to be recognized. If the program fails to detect a picture automatically, use the **Picture** tool to draw a picture area around the picture manually.

1. On the toolbar at the top of the **Image** window, select  tool and then use the mouse to draw a rectangle around the picture on your image.
2. Run the OCR process gain.

If there is text printed over a picture in your document, draw a **Background Picture** area around the picture.

1. In the **Image** window, select the  tool on the toolbar and draw a frame around the picture.
Note: If you do not see this tool on the toolbar, add it. See Toolbars for detailed instructions.
2. Start the OCR process.

You can change an existing area to **Picture** or **Background Picture**. Right-click the area and select **Change Area Type>Picture** or **>Background Picture** respectively.

Barcode Not Detected

Barcode areas mark the barcodes contained in a document. If your document contains a barcode that you would like to render as a sequence of letters and digits rather than as an image, select the barcode and change the area type to **Barcode**.

Note: Barcode recognition is disabled by default. To enable barcode recognition, select **Tools>Options...**, click the **Read** tab, and select **Look for barcodes**.

To draw a barcode area:

1. Hold **CTRL+SHIFT+ALT** and use the mouse to select a barcode in the **Image** window or in the **Zoom** window.
2. Run the OCR process.



Code 3 of 9	Code 128	Check Interleaved 2 of 5	UPC-A	QR code
Check Code 3 of 9	EAN 8	Matrix 2 of 5	UPC-E	Patch code
Code 3 of 9 without asterisk	EAN 13	Postnet	PDF417	
Codabar	IATA 2 of 5	Industrial 2 of 5	Aztec code	
Code 93	Interleaved 2	UCC-128	Data	

Note: The default barcode type is **Autodetect**, which means that barcode types are to be detected automatically. However, not all barcode types can be autodetected, and sometimes you may need to select the appropriate type manually on the **Area Properties** panel.

ABBYY FineReader can automatically detect the following types of barcode: Code 3 of 9, Code 93, Code 128, EAN 8, EAN 13, Postnet, UCC-128, UPC-E, PDF417, UPC-A, QR code.

Table Not Detected

Table areas mark tables or text fragments arranged in a tabular form. If the program fails to detect a table automatically, use the **Table** tool to draw a table area around the table.

1. Select the  tool on the toolbar at the top of the **Image** window and use the mouse to draw a rectangle around the table on your image.
2. Right-click the table and select **Analyze Table Structure** from the shortcut menu or click the  button on the pop-up toolbar of the **Table** area.
3. Use the pop-up toolbar, the table shortcut menu, or the commands from the main **Areas** menu to edit the results.
4. Run the OCR process again.

You can change the type of an area that has already been recognized. To change the type of an area to **Table**, right-click the area and select **Change Area Type>Table** from the shortcut menu.

Table Cells Detected Incorrectly

If ABBYY FineReader failed to correctly detect the rows and columns in a table automatically, you can adjust the layout manually and run the OCR process again.


[illegible]

Editing tables in the Image window or in the Zoom window

- Use the commands on the **Areas** menu or the commands on the shortcut menu to:
 - **Analyze Table Structure** (automatically analyzes the structure of the table and saves you the trouble of drawing vertical and horizontal separators manually);
 - **Split Cells**, **Merge Cells**, and **Merge Rows** (allows you to correct any wrongly identified rows or columns)

- Use the pop-up toolbar of the selected table to:

Add a horizontal or vertical separator, delete table separators, analyze table structure, split table cells, or merge table cells or rows (click the **Table** area and select the appropriate tool from the pop-up toolbar: ).

Tip. Click the  tool on the pop-up toolbar to select objects.

Notes:

1. The cell splitting command can only be applied to table cells that were previously merged.
2. To merge or split table cells or rows, select them on the image and then choose the appropriate tool or menu command.
3. If the table contains vertical text cells, you may need to change the text properties first.
4. If the table cells only contain an image, check the **Treat selected cells as picture** option on the **Area Properties** panel (the **>Area Properties** shortcut menu).
If the cell contains a combination of image and text, select the image to turn it into a dedicated area inside the cell.

Adjusting Text Area Properties

When performing automatic document analysis, ABBYY FineReader detects **different area types** on document pages (such as tables, images, text and barcodes). The program also analyzes the document structure and identifies headings and other layout elements such as headers and footers, images and diagram captions, etc.



You can manually change incorrectly defined properties in some **Text** and **Table** areas.

1. Select the area in the **Image** window or the **Zoom** window and adjust the properties in the **Area Properties** panel of the **Image** window.

The **Area Properties** panel lets you set the following properties:

- **Language** — For details, see Document languages.
- **Function** displays the function of the text, i.e. header or footer, heading, etc. (applies only to **Text** areas).
- **Orientation** — For details, see Vertical or Horizontal Text Not Recognized Properly.
- **Print type** — For details, see Print Type.
- **Inversion** — For details, see Vertical or Horizontal Text Not Recognized Properly.
- **Direction of hieroglyphic text** — For details, see Recognizing Documents in More Than One Language.

Notes:

- a. Use / buttons to navigate left and right along the properties panel if the **Image** window is not wide enough to display the entire properties panel.
 - b. Some of the text properties can be modified using the shortcut menu that becomes available when you right-click a **Text** area.
2. Run the OCR process again after making the necessary changes.

Vertical or Inverted Text Not Recognized Properly

A fragment of recognized text may contain a large number of errors if the orientation of the fragment was detected incorrectly or if the text in the fragment is inverted (i.e. light text is printed on a dark background).

To solve this problem:

1. In the **Image** window, select the area or the table cell that contains vertical or inverted text.
2. In the **Image** window on the **Area Properties** panel (right-click the area and select **Area Properties** to make the panel visible), select:

Horizontal Script or Vertical Script from the Direction of hieroglyphic text drop-down list

or

Inverted from the **Inversion** drop-down list
3. Run the OCR process again.

Paper Document Contains Decorative (Non-Standard) Fonts

The Training mode improves OCR quality on documents with decorative fonts or documents containing special characters (e.g. mathematical symbols).

In Training mode, a user pattern is created which can be used when performing OCR on the entire text.

To use a **pattern** to recognize a document:

1. Open the **Options** dialog box (**Tools>Options...**) and then click the **Read** tab.
2. Under Training, select the Use only user pattern option.

Note: If you select **Use built-in and user patterns**, ABBYY FineReader will use both the user patterns and its factory preset patterns for OCR.
3. Click the **Pattern Editor...** button.
4. In the **Pattern Editor** dialog box, select the desired pattern and then click **OK**.
5. In the ABBYY FineReader main window, click the **Read** button.

Incorrect Font in Recognized Text or Some Characters Are Replaced with "?" or "□"

If you see "?" or "□" replacing letters in the **Text** window, check to see if the fonts selected to display the recognized text contain all the characters used in your text.

You can change the font in the recognized text without launching the OCR process again.

To change the font in a short document:

1. Select a text fragment where some characters are displayed incorrectly.
2. Right-click the selection and select **Properties** from the shortcut menu.
3. On the **Text Properties** panel, select the desired font. The font in the selected text fragment will change accordingly.

To change the font in a long document where styles are used:

1. On the **Tools** menu, click **Style Editor...**
2. In the **Style Editor** dialog box, select the desired style and change its font.
3. Click **OK**. The font in all the text fragments using this style will change accordingly.

Important! If an ABBYY FineReader document was recognized or edited on another computer, the text in the document may be displayed incorrectly on your computer. If this is the case, make sure that all the fonts used in this document are installed on your machine.

Checking and Editing the Recognized Text




Once the OCR process is complete, the recognized text appears in the **Text** window for review and editing. This section provides you with information on working with recognized text.

- Checking the Text in the Text Window
- Checking Spelling
- User Dictionary: Adding and Removing Words
- Using Styles
- Editing Out Confidential Information
- Editing Hyperlinks
- Editing Tables

Checking Text in the Text Window

You can check, edit, and format recognition results in the **Text** window.

The **Text** window toolbar in the upper part of the window lets you open the **Verification** dialog box. This dialog box activates spellchecking, which can also be run by using the spellchecker control buttons in the **Text** window.

Use the   buttons to go to the next/previous uncertain word or character. If the uncertain characters are not highlighted, click the  button on the toolbar of the **Text** window.

To check an uncertain word in the **Text** window:

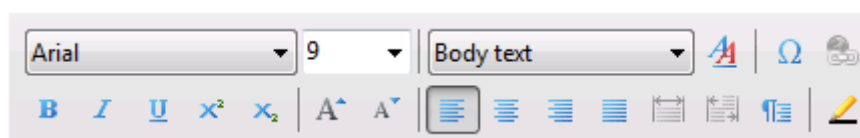
1. Click the word in the **Text** window. The word will be highlighted in the **Image** window, and in the **Zoom** window you will see its magnified image.
2. Make changes in the word in the **Text** window, if necessary.

This method is convenient when you need to compare the source and resulting documents.



The Main Toolbar provides access to the drop-down saving options menu.

The tools for formatting recognized text are available on:

- **The right-hand side** of the Main Toolbar.



- The **Text Properties** panel (right-click and select **Text>Properties**).

Note: Use the / buttons to navigate left and right along the properties panel if the **Text** window is not wide enough to display the entire text properties panel.

Checking the spelling in the Verification dialog box

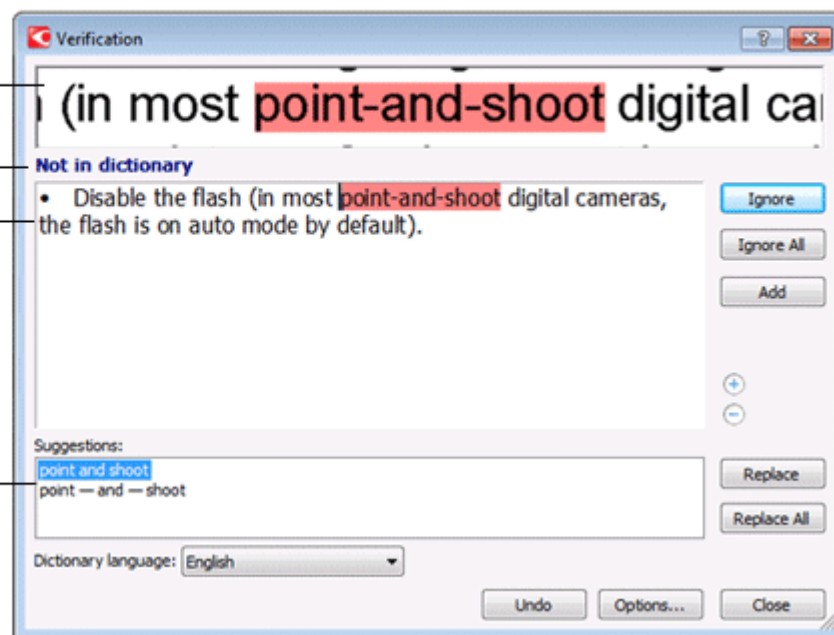
You can verify words with low-confidence characters in the **Verification** dialog box (**Tools>Verification...**).

Identical to the Zoom window,
show words with uncertain characters

Type of error

Recognized text

Recognition variants

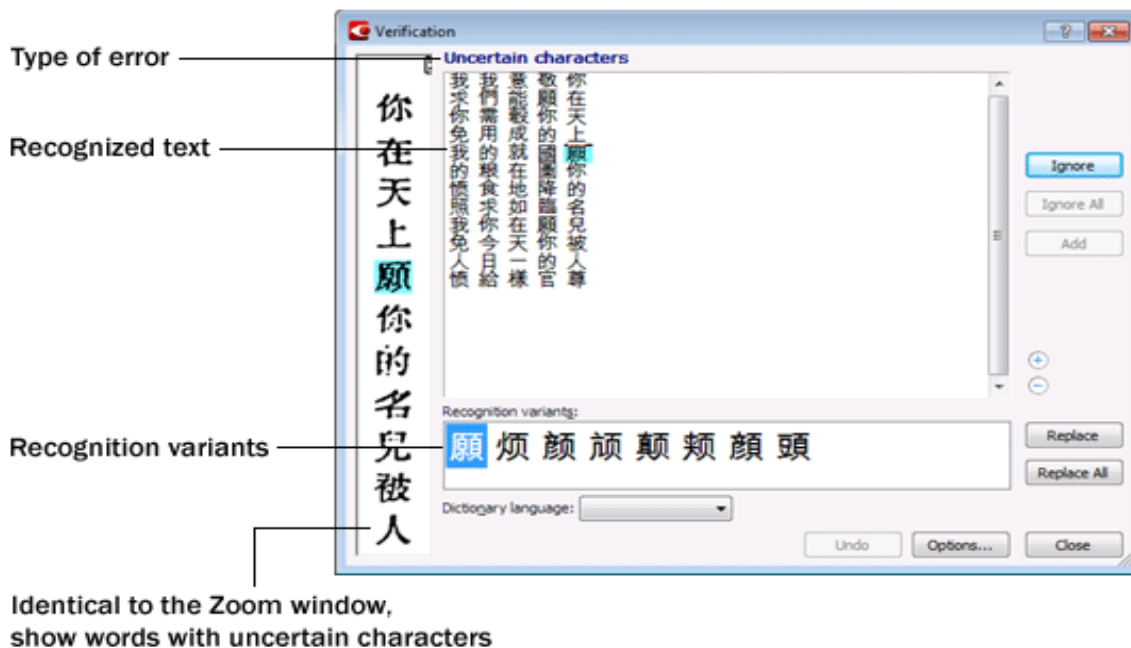


In this dialog box, you can choose to apply (or reject) the suggested changes either to the current occurrence or to all the occurrences of a word in the text.

- Click **Ignore** to skip the word without making changes.
The highlighting will be removed.
- Click **Replace** to replace the word with one of the suggested corrections.
- Click **Add** to place the selected word into the dictionary.
This word will no longer be considered an error.

Tip. To change the spelling check options, click **Options...**

For character-based languages, the program will suggest similar-looking characters when no direct matches are found.



User Dictionary: Adding and Removing Words

If your text contains many special terms, abbreviations, and proper names, adding these terms to the user dictionary will improve OCR quality. For each added word, ABBYY FineReader generates all possible word forms to allow the program to reliably recognize it in its various forms (plural, verb, etc.).

To add a word to the dictionary while checking the spelling:

Select **Tools>Verification...** and click the **Add** button in the **Verification** dialog box

or

In the **Text** window, right-click the word and select the **Add to Dictionary** command

You can also add and remove words while viewing a user dictionary for a particular language.

1. On the **Tools** menu, click **View Dictionaries...**
2. In the **User Dictionaries** dialog box, select the desired language and click **View...**
3. In the **Dictionary** dialog box, type a word and click **Add** or select a word and click **Delete**.

The program will notify you if the desired word is already in the dictionary.

You can import user dictionaries created in previous versions of ABBYY FineReader (versions 8.0, 9.0, and 10 are supported).

1. On the **Tools** menu, click **View Dictionaries...**, select the desired language, and click **View...**
2. In the **Dictionary** dialog box, click the **Import...** button and select the dictionary file you wish to import (it must have the extension *.pmd, *.txt, or *.dic).

Using Styles

When recognizing a text, ABBYY FineReader detects the styles and formatting used in the original document. These styles and formatting are reproduced in the output document. You can review and edit the document styles and create new styles to format recognized text in the **Text** window.

To apply a style to a selected text fragment:

1. Select the desired text fragment in the **Text** window.
2. Select **Properties** from the shortcut menu.
3. Select the desired style in the **Text Properties** panel from the **Style** list.

Note. When saving recognized texts in RTF, DOC, DOCX, and ODT formats, all styles are preserved.


Changing, creating, and merging styles:

1. On the **Tools** menu, click **Style Editor...**
2. In the **Style Editor** dialog box, select the desired style and adjust its name, font, font size, character spacing, and scale.
3. To create a new style, click **New**. The newly created style will be added to the list of existing styles where you can adjust it.
4. To merge multiple styles into one, select the styles to merge and click **Merge....** In the **Merge Styles** dialog box, specify the style into which to merge the selected styles.
5. Click **Save** to save the changes.

You can navigate between text fragments printed in identical styles. In the **Style Editor**, select the desired style and click **Previous Fragment** or **Next Fragment**.

Editing Out Confidential Information (ABBYY FineReader Corporate Edition only)

In ABBYY FineReader 11, you can easily remove confidential information from a recognized text

1. On the **Tools** menu, click **Redaction Mode** or click the  button on the main toolbar. The mouse pointer will change to a marker.
2. In the **Text** window, use the marker to black out the text you wish to conceal.


Tip. If you black out some characters by mistake, you can undo the last redaction by pressing **CTRL+Z** or clicking the **Undo** on the main toolbar.

3. Save your document.

The redacted text will appear as dots in the output document. If the saving format you selected supports text and background colors, black dots will be printed against a black background.

Note: When you save a page, the redacted areas will appear as black rectangles.

To switch off the **Redaction** mode, either


- Select **Tools>Redaction Mode** once again or
- Click the  button on the main toolbar

Editing Hyperlinks

ABBYY FineReader detects hyperlinks and recreates their destination addresses in the output document. Detected hyperlinks are underlined and displayed in blue.

When viewing the recognized document in the **Text** window, rest the mouse pointer on a hyperlink to view its address. To follow a hyperlink, select **Open Hyperlink** from its shortcut menu, or press CTRL and left-click the hyperlink.

To add, delete or change the text or address of a hyperlink:

1. In the **Text** window, select the desired hyperlink.
2. To remove a hyperlink, right-click it and select **Remove Hyperlink** from the shortcut menu.
3. To add or change a hyperlink, click **Hyperlink...** in its shortcut menu, or click  on the main toolbar at the top of the **Text** window. In the **Edit Hyperlink** dialog box, you can:
 - a. Make the necessary text changes in the **Text to display** field.
 - b. Select/change the hyperlink type in the **Link to** group:
 - Select **Web page** to link to an Internet page.
In the **Address** field, specify the protocol and the URL of the page (e.g. *http://www.abbyy.com*).
 - Select **Local file** to link to a file.
Click **Browse...** to browse for the file to which the hyperlink will point (e.g. *file:///D:/MyDocuments/ABBYY FineReaderGuide.pdf*).
 - Select **E-mail address** so that the user can send an e-mail message to the address contained in the hyperlink by simply clicking the hyperlink.
In the **Address** field, specify the protocol and the e-mail address (e.g. *mailto:office@abbyy.com*).

Editing Tables

ABBYY FineReader lets you edit recognized tables in the **Text** window. The following options are available:

1. Split table cells.

Click the left mouse button to select a cell, and then select **Split Table Cells** from the **Edit** menu.

Important! This command can only be applied to table cells that have been previously merged.

2. Merge table cells.

Use the mouse to select the table cells to be merged, and then select **Merge Table Cells** from the **Edit** menu.

3. Merge table rows.

Use the mouse to select the table rows to be merged, and then select **Merge Table Rows** from the **Edit** menu.

4. Delete cell contents.

Select the cell (or a group of cells) with the contents you wish to delete and press the DELETE key.

Note: By default, the table editing tools are not displayed on the toolbar. You can add buttons to the toolbar by using the **Customize Toolbars and Shortcuts** dialog box (**Tools>Customize...**).

Working with Complex-Script Languages

With ABBYY FineReader, you can also recognize documents in Hebrew, Yiddish, Japanese, Chinese, Thai, Korean, and Arabic languages. Consider the following when working with documents in Chinese, Japanese or Korean languages and documents in which a combination of character-based and European languages is used.

You may need to do the following to recognize these types of documents:

- Installing Additional Languages
- Recommended Fonts

This section contains some tips on how to improve OCR quality:

- Disabling Automatic Image Processing
- Recognizing Documents Written in More Than One Language
- Non-European Characters Not Displayed in the Text Window
- Selecting the Direction of Recognized Text

Installing Additional Languages

To be able to recognize texts written in Japanese, Chinese, Thai, Korean, Arabic, Hebrew or Yiddish, you may need to install these languages.

Note: Microsoft Windows Vista and Windows 7 support these languages by default.

To install new languages in Microsoft Windows XP:

1. Click Start on the Control panel.
2. Select Control Panel>Regional and Language Options.
3. On the **Language** tab, select:
 - **Install files for complex-script and right-to-left languages**
to be able to recognize texts in Hebrew, Yiddish, Arabic, and Thai
 - **Install files for East Asian languages**
to be able to recognize texts in Japanese, Chinese, and Korean
4. Click **OK**.

Recommended Fonts

The table below lists the recommended fonts for working with Hebrew, Yiddish, Thai, Chinese, and Japanese texts.

OCR Language	Recommended Font
Arabic	Arial™ Unicode™ MS*
Hebrew	Arial™ Unicode™ MS*
Yiddish	Arial™ Unicode™ MS*
Thai	Arial™ Unicode™ MS* Aharoni David Levenim mt Miriam Narkisim Rod
Chinese (Simplified), Chinese (Traditional), Japanese, Korean, Korean (Hangul)	Arial™ Unicode™ MS* SimSun fonts For example: SimSun (Founder Extended), SimSun–18030, NSimSun. Simhei YouYuan PMingLiU MingLiU Ming(for–ISO10646) STSong

* This font is installed together with Microsoft Windows XP and Microsoft Office 2000 or later.

Disabling Automatic Image Processing

By default, any pages you add to an ABBYY FineReader document are automatically recognized.

However, if your document contains a text in a Character-based language combined with a European language, we recommend disabling automatic page orientation detection and using the dual page splitting option only if all of the page images have the correct orientation (e.g., they were not scanned upside down).

The **Detect page orientation** and **Split facing pages** options can be enabled and disabled directly in the image scanning and opening dialog boxes, and from the **Options** dialog box on the **Scan/Open** tab.

Note: To split facing pages in Arabic, Hebrew, or Yiddish, be sure to select the corresponding recognition language first and only then select the **Split facing pages** option. This will ensure that the pages are arranged in the correct order. You can also restore the original page numbering by selecting the **Swap book pages** option. For details, see *Numbering Pages in an ABBYY FineReader Document*.

If your document has a complex structure, we recommend disabling automatic analysis and OCR for images and performing these operations manually.

To disable automatic analysis and OCR:

1. Open the **Options** dialog box (**Tools>Options...**).
2. Select the **Do not read and analyze acquired page images automatically** option on the **Scan/Open** tab.
3. Click **OK**.


Recognizing Documents Written in More Than One Language

In the instructions below, a document that contains both English and Chinese text is taken as an example.

1. Disable the automatic analysis and OCR options.
2. On the main toolbar, select **More languages...** from the **Document Languages** drop-down list. Select **Specify languages manually** from the **Language Editor** dialog box and select Chinese and English from the language list (for details, see *Document Languages*).
3. Scan or open images after disabling **Detect page orientation**. The dual page splitting option should be used only if all page images have the correct orientation. The pages will be added to the current ABBYY FineReader document after the command is executed.

Important! When scanning, be sure that the pages are properly centered on the scanner's glass bed. If the skew is too large, the text may be recognized incorrectly.

4. To draw areas on the image manually, use the tools for **Adjusting Area Shapes and Area Borders**.

Note: If the structure of your document is simple, you can launch automatic layout analysis. Click the  (**Analyze**) button on the toolbar of the **Image** window or press CTRL+E.

5. If there are areas on the image where text is written in only one language:
 - a. Select these areas.
 - b. Select the language of the text area (Chinese or English) on the **Area Properties** panel.

Important! You can only specify a language for areas of the same type. If you select both **text** and **table** areas, you won't be able to specify a language.

- c. If necessary, select the text direction from the **Orientation** drop-down menu (for details, see *Vertical or Inverted Text Not Recognized Properly*).
 - d. For texts in character-based languages, the program provides a selection of the text directions in the **Direction of hieroglyphic text** drop-down menu (for details, see *Changing Text Properties*).
6. Click **Recognize**.

Non-European Characters Not Displayed in the Text Window

If text in a character-based language is displayed incorrectly in the **Text** window, you may have selected the **Plain text** mode.


To change the font used in **Plain text** mode:

1. Open the **Options** dialog box (**Tools>Options...**).
2. Go to the **View** tab.
3. Select Arial Unicode MS from the Font used to display plain text drop-down menu.
4. Click **OK**.

If nothing has changed in the **Text** window, refer to: Incorrect Font in Recognized Text or Some Characters Are Replaced with "?" or "□".

Selecting the Direction of Recognized Text

ABBYY FineReader automatically detects text direction when it performs OCR. If required, you can manually adjust the direction of recognized text.

1. Go to the **Text** window.
2. Select one or several paragraphs.
3. Click  on the main toolbar.

Note: For character-based languages, use the **Direction of hieroglyphic text** option to select the text direction before text recognition is performed. For details, see Changing Text Properties.

Saving the Results

Recognized texts can be saved to a file, sent to another application without saving them to disk, copied to the Clipboard, or sent by e-mail as attachments in any of the supported saving formats.

- **Saving: General**
Describes the saving capabilities provided by ABBYY FineReader.
- **Document Properties**
- **Saving text in RTF, DOC, DOCX or ODT format**
- **Saving in XLSX**
- **Saving in PDF**
- **Saving in PDF/A**
- **PDF Security Settings**
Explains the security settings available when saving in PDF: protecting your document with passwords that prevent unauthorized opening, editing, or printing and selecting an encryption level compatible with earlier versions of Adobe Acrobat.
- **Saving in HTML**
- **Saving in PPTX**
- **Saving in TXT**

- Saving in CSV
- Saving E-books
- Saving in DjVu
- Saving to Microsoft SharePoint
- Saving an Image of the Page
Describes the procedure that saves your page without performing OCR on it and provides advice on reducing the size of your images.

Saving: General

The **File** menu offers you a choice of different saving methods for the recognized text. You can also send the recognized text to various applications.

- **File>Save FineReader Document**
Saves the current ABBYY FineReader document. Both the recognized text and the page images are saved.
- **File>Save Document As**
Saves the recognized text on your hard disk in a format of your choice.
- **File>Send Document To**
Opens the recognized text in an application of your choice. No information is saved on your drive.
- **File>Save To Microsoft SharePoint (Corporate Edition only)**
Saves the recognized text in a network location: on a website, on an intranet portal, or in an electronic library.
- **File>E-mail**
Sends the image or recognized text via e-mail. In the dialog box that opens, select the desired options for your e-mail attachment and click **OK**. A new e-mail message will be created with the image or recognized text attached to it.
- **File>Print**
Prints the text or the images of the selected pages of the current ABBYY FineReader document.

Supported applications

- Microsoft Word 2000 (9.0), 2002 (10.0), 2003 (11.0), 2007 (12.0), and 2010 (14.0)
- Microsoft Excel 2000 (9.0), 2002 (10.0), 2003 (11.0), 2007 (12.0), and 2010 (14.0)
- Microsoft PowerPoint 2003 (11.0) (with Microsoft Office Compatibility Pack for Word, Excel, and PowerPoint 2007 formats), 2007 (12.0), and 2010 (14.0)
- Corel WordPerfect 10.0 (2002), 11.0 (2003), 12.0, 13.0, and 14.0
- Lotus Word Pro 97 and Millennium Edition
- OpenOffice.org 3.0, 3.1
- Adobe Acrobat/Reader (5.0 and later)

Note: To ensure better compatibility, we recommend installing the latest updates and upgrades available for the above applications.

Document Properties

Document properties contain information about the document (the extended title of the document, author, subject, key words, etc). Document properties can be used to sort your files. Additionally, you can search for documents by their properties.

When recognizing PDF documents and certain types of image files, ABBYY FineReader will export the properties of the source document. You can then edit these properties.

To add or modify document properties:

1. Click Tools>Options...
2. Click the **Document** tab, and, in the **Document properties group**, specify the title, author, subject and key words.

Saving text in RTF, DOC, DOCX or ODT format

To save your text in RTF/DOC/DOCX/ODT:

- In the drop-down list on the main toolbar, choose a document layout saving mode.
- Click **File>Save Document As>Microsoft Word 97–2003 Document** (to save to ODT format, choose **File>Save Document As>OpenOffice.org Writer Document**) or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **RTF/DOCX/ODT** tab.

The saving options on this tab are grouped into the following categories:

Retain layout

Depending on how you are planning to use your electronic document, select the best option below:

- **Exact copy**
Produces a document that maintains the formatting of the original. This option is recommended for documents with complex layouts, such as promotion booklets. Note, however, that this option limits the ability to change the text and formatting of the output document.
- **Editable copy**
Produces a document that nearly preserves the original format and text flow but allows easy editing.
- **Formatted text**
Retains fonts, font sizes, and paragraphs, but does not retain the exact locations of the objects on the page or the spacing. The resulting text will be left-aligned (right-to-left texts will be right-aligned).

Note: Vertical texts will be changed to horizontal in this mode.
- **Plain text**
Unlike the **Formatted text** mode, this mode does not retain formatting

Default paper size

You can select the paper size to be used for saving in RTF, DOC, DOCX, or ODT format from the **Default paper size** drop-down list.

Tip. To ensure the recognized text fits the paper size, select the **Increase paper size to fit content** option. ABBYY FineReader will automatically select the most suitable paper size when saving.

Text settings

- **Keep headers and footers**
Retains running titles (headers and footers) in the output text.
- **Keep page breaks**
Retains the original page arrangement.
- **Keep line breaks**
Retains the original arrangement into lines.
- **Keep line numbers**
Retains the original line numbering (if any). The line numbers will be saved in a separate field that remains unchanged when you edit the text.
Note: This feature is only available if **Exact copy** or **Editable copy** is selected.
- **Keep text and background colors**
Retains the original color of the letters.

Note: Word 6.0, 7.0, and 97 (8.0) have a limited text and background color palette, therefore the original document colors may be replaced with the ones available in the Word palette. Word 2000 (9.0) or later retains the colors of the source document in full.

Picture settings

Documents containing a large number of pictures are very large. To reduce the size of the file, select the desired option in the **Image quality** group.

Tip:

- To change the picture saving parameters, click **Custom....** In the **Custom Settings** dialog box, select the desired parameters and click **OK**.
- If you don't want to keep pictures in the recognized text, make sure the **Keep pictures** option is clear.

Advanced

Some of the more advanced saving options become available by clicking the **Advanced** group.

- **Highlight uncertain characters**
Select this option to edit the recognized text in Microsoft Word rather than in the ABBYY FineReader **Text** window. All uncertain characters will be highlighted in the Microsoft Word window.

Tip. You can change the color of uncertain characters on the **View** tab of the **Options** dialog box (**Tools>Options...**).

- **Enable compatibility with other word processors**
Produces a document that can be opened and edited in earlier versions of Microsoft Word and other word processing applications that support the RTF format.

Saving in XLSX

To save your text in XLS/XLSX:

- Click **File>Save Document As>Microsoft Excel 97–2003 Workbook** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **XLSX** tab.

The following options are available:

- **Ignore text outside tables**
Saves only the tables and ignores the rest.
- **Convert numeric values to numbers**
Converts numbers into the "Numbers" format in the XLS file. Microsoft Excel may perform arithmetical operations on cells of this format.
- **Keep headers and footers**
Preserves headers and footers in the output document.

Saving in PDF

To save your text in PDF:

- Click **File>Save Document As>PDF Document** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **PDF** tab.

The saving options on this tab are grouped into the following categories:

Default paper size

From the **Default paper size** drop-down list, choose the paper size to be used for saving in PDF format.

Save mode

Depending on how you are planning to use your electronic document, select one of the best option below:

- **Text and pictures only**
This option saves only the recognized text and the associated pictures. The page will be fully searchable and the size of the PDF file will be small. The appearance of the resulting document may slightly differ from the original.
- **Text over the page image**
This option saves the background and pictures of the original document and places the recognized text over them. Usually, this PDF type requires more disk space than **Text and pictures only**. The resulting PDF document is fully searchable. In some cases, the appearance of the resulting document may slightly differ from the original.
- **Text under the page image**
This option saves the entire page image as a picture and places the recognized text

underneath. Use this option to create a fully searchable document that looks virtually the same as the original.

- **Page image only**

This option saves the exact image of the page. This type of PDF document will be virtually indistinguishable from the original but the file will not be searchable.

Depending on the save mode you select, some of the following options will become available:

- **Keep text and background colors**

Select this option to retain the font color and background when saving in PDF.

- **Keep headers and footers**

Preserves headers and footers in the output document.

- **Enable Tagged PDF (compatible with Adobe Acrobat 5.0 and above)**

Select this option to add PDF tags to the output PDF document.

Apart from text and pictures, PDF files can contain information about the document structure such as logical parts, pictures, and tables. This information is encoded in PDF tags. A PDF file equipped with PDF tags may be reflowed to fit different screen sizes and will display well on handheld devices.

Important! If your recognized document contains text in Hebrew or Yiddish, you cannot create a tagged PDF file.

- **Use Mixed Raster Content**

Select this option to retain the visual quality of the text and the pictures in a highly compressed document.

- **Create outline**

Select this option if you wish to create a table of contents from the document headings.

Image settings

The resulting file may be quite large if your text contains a lot of pictures or if you choose to save the page image together with the recognized text. To adjust the size of the resulting file and the quality of the pictures, select one of the options from the **Image settings** drop-down list:

- **Best quality (source image resolution)**

Select this option to retain the quality of the pictures or the page image. The original resolution of the source image will be preserved.

- **Balanced**

Select this option to reduce the size of the PDF file while keeping the quality of the pictures or page image at a reasonably high level.

- **Compact size**

Select this option to obtain a small-sized PDF file. The resolution of the pictures and page image will be reduced to 150 dpi, affecting their quality.

- **Custom...**

Select this option to make your own picture and page image settings in the **Custom Settings** dialog box.

Tip. If you do not want to save pictures in the recognized text, make sure the **Keep pictures** option is not selected.

Font settings

When saving texts to PDF, you can use either Adobe fonts or Windows fonts installed on your machine. To specify the font set to be used, in the **Font** drop-down list, select one of the following:

- **Use predefined fonts**
The PDF file will use Adobe fonts, for example: Times New Roman, Arial, Courier New.
- **Use Windows fonts**
The PDF file will use the Windows fonts installed on your computer.

If you do not want to use embedded fonts in your PDF document, make sure that the **Embed fonts** option is not selected.

Security

You can use passwords to prevent unauthorized opening, printing or editing of your PDF document:

- Click the **PDF Security Settings...** button and in the dialog box, select the desired security settings.

Saving in PDF/A

Select the PDF/A saving format if you want to create a searchable PDF document that looks like the original or if you need to store your document in an archive.

To save your text in PDF/A:

- Click **File>Save Document As>PDF/A Document** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **PDF/A** tab.

The saving options on this tab are grouped into the following categories:

Default paper size

From the **Default paper size** drop-down list, choose the paper size to be used for saving in PDF format.

Save mode

Depending on how you are planning to use your electronic document, select one of the best option below:

- **Text and pictures only**
This option saves only the recognized text and the associated pictures. The page will be fully searchable and the size of the PDF file will be small. The appearance of the resulting document may slightly differ from the original.
- **Text over the page image**
This option saves the background and pictures of the original document and places the recognized text over them. Usually, this PDF type requires more disk space than **Text and pictures only**. The resulting PDF document is fully searchable. In some cases, the appearance of the resulting document may slightly differ from the original.
- **Text under the page image**
This option saves the entire page image as a picture and places the recognized text underneath. Use this option to create a fully searchable document that looks virtually the same as the original.

- **Page image only**

This option saves the exact image of the page. This type of PDF document will be virtually indistinguishable from the original but the file will not be searchable.

Depending on the save mode you select, some of the following options will become available:

- **Keep text and background colors**

Select this option to retain the font color and background when saving in PDF.

- **Keep headers and footers**

Preserves headers and footers in the output document.

- **Enable Tagged PDF (compatible with Adobe Acrobat 5.0 and above)**

Select this option to add PDF tags to the output PDF document.

Apart from text and pictures, PDF files can contain information about the document structure such as logical parts, pictures, and tables. This information is encoded in PDF tags. A PDF file equipped with PDF tags may be reflowed to fit different screen sizes and will display well on handheld devices.

Important! If your recognized document contains text in Hebrew or Yiddish, you cannot create a tagged PDF file.

- **Use Mixed Raster Content**

Select this option to retain the visual quality of the text and the pictures in a highly compressed document.

- **Create outline**

Select this option if you wish to create a table of contents from the document headings.

Picture settings

The resulting file may be quite large if your text contains a lot of pictures or if you choose to save the page image together with the recognized text. To adjust the size of the resulting file and the quality of the pictures, select one of the options from the **Image settings** drop-down list:

- **Best quality (source image resolution)**

Select this option to retain the quality of the pictures or the page image. The original resolution of the source image will be preserved.

- **Balanced**

Select this option to reduce the size of the PDF file while keeping the quality of the pictures or page image at a reasonably high level.

- **Compact size**

Select this option to obtain a small-sized PDF file. The resolution of the pictures and page image will be reduced to 150 dpi, affecting their quality.

- **Custom...**

Select this option to make your own picture and page image settings in the **Custom Settings** dialog box.

Tip. If you do not want to save pictures in the recognized text, make sure the **Keep pictures** option is not selected.

PDF Security Settings

When you save a recognized text in PDF format, you can protect the contents of the PDF–document with:

1. a password required to open the document
2. a password required to modify or print the document or copy information to another application
3. encoding


Passwords and other PDF security parameters are set in the **PDF Security Settings** dialog box.

To open it, click the **PDF Security Settings...** button located on the bottom of the **Save>PDF** tab of the **Options** dialog box (**Tools>Options...**).

Password for opening the document

This password protects the PDF–document from unauthorized opening. The user will be able to open the document only after entering the password specified by the author.

To set a password required to open the document:


1. Select **Require password to open document**.
2. Click .
3. In the **Enter Document Open Password** dialogue box, enter your password and confirm it.

The password you specified will be displayed as dots in the **Document Open Password** field of the **PDF Security Settings** dialogue box.

Password for editing and printing the document

This password protects the PDF–document from unauthorized editing and printing as well as copying information to another application. The user will be able to do the described action only after entering the password specified by the author.

To set a password required to access file contents:

1. Select **Restrict printing and editing the document and its security settings**.
2. Click .
3. In the **Enter Permissions Password** dialogue box, enter your password and confirm it.

The password you specified will be displayed as dots in the **Permissions Password** field of the **PDF Security Settings** dialogue box.

Next, select the actions you want to allow the users to do with the PDF–document. These restrictions are set in the **Permissions settings** group.

- Items in the **Printing allowed** drop–down list let you allow or prohibit document printing.
- Items in the **Changes allowed** drop–down list let you allow or prohibit document editing.

- If the **Enable copying text, pictures, and other content** option is selected, the user will be able to copy the contents of the PDF–document (text and images) to Clipboard. If you want to prohibit such actions, make sure this option is clear.
- If the **Enable text access for screen reader devices** option is selected, the user will be able to make screenshots of the opened PDF–document. If you want to prohibit such actions, make sure this option is clear. Note that this option is only available for PDF–documents with 128–bit encoding or higher.

Encoding level

Items in the **Encryption level** drop–down list let you set the encoding type for the password–protected PDF–document.

Select one of the encoding levels:

- **Low (40 bit) – compatible with Adobe Acrobat 3.0 and above** sets a low 40–bit encoding level based on the RC4 standard.
- **High (128 bit) – compatible with Adobe Acrobat 5.0 and above** sets a high 128–bit encoding level based on the RC4 standard, but the users of earlier Adobe Acrobat versions will not be able to open such PDF–document.
- **High (128–bit AES) – compatible with Adobe Acrobat 7.0 and above** sets a high 128–bit encoding level based on the AES standard, but the users of earlier Adobe Acrobat versions will not be able to open such PDF–document.

Saving in HTML

To save your text in HTML:

- In the drop–down list on the main toolbar, choose a document layout saving mode.
- Click **File>Save Document As>HTML Document** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **HTML** tab.

The saving options on this tab are grouped into the following categories:

Retain layout

Depending on how you are planning to use your electronic document, select the appropriate option:

- **Flexible layout**
Produces a document that maintains the formatting of the original. The output document can be easily edited.
- **Formatted text**
Retains fonts, font sizes, and paragraphs, but does not retain the exact locations of the objects on the page or the spacing. The resulting text will be left–aligned. (Right–to–left texts will be right–aligned.)

Note: Vertical texts will be changed to horizontal in this mode.

- **Plain text**

This mode does not retain formatting.

Use CSS

Select this option to save the file in HTML 4 format, which allows maintaining the document layout by using a style table incorporated in the HTML-file.

Save mode

- **Keep line breaks**

Retains the original arrangement of lines.

- **Retain text and background colors**

Retains the original color of the letters.

- **Keep headers and footers**

Preserves headers and footers in the output document.

Picture settings

Documents containing a large number of pictures are very large. To reduce the size of the file, select the desired option in the **Image quality** group.

Tip:

- To change the picture saving parameters, click **Custom....** In the **Custom Settings** dialog box, select the desired parameters and click **OK**.
- If you don't want to keep pictures in the recognized text, make sure the **Keep pictures** option is clear.

Character encoding

ABBYY FineReader detects the code page automatically. To change the code page or its type, select the desired code page or code page type from the drop-down lists under **Character encoding**.

Book settings

Adjust **Book settings** if you are converting a printed book into electronic form. ABBYY FineReader can automatically save book chapters into separate HTML files and recreate links to them in the table of contents. You can also split the document into separate HTML files based on level 1 or level 2 headings.

Saving in PPTX

To save your text in PPTX:

- Click **File>Save Document As>Microsoft PowerPoint Presentation** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **PPTX** tab.

The saving options on this tab are grouped into the following categories:

Text settings

- **Keep line breaks**
Retains the original arrangement of lines.
- **Wrap text**
All recognized text is fit entirely inside the text block in the slide.
- **Keep headers and footers**
Preserves headers and footers in the output document.

Picture settings

Documents containing a large number of pictures are very large. To reduce the size of the file, select the desired option in the **Image quality** group.

Tip:

- To change the picture saving parameters, click **Custom....** In the **Custom Settings** dialog box, select the desired parameters and click **OK**.
- If you don't want to keep pictures in the recognized text, make sure the **Keep pictures** option is clear.

Saving in TXT

To save your text in TXT:

- Click **File>Save Document As>Text Document** or the **Save** button on the main toolbar. Click the arrow next to the **Save** button and choose a saving format from the list. If there is no suitable format in the list, click **Save to Other Formats...**, and, in the dialog box that opens, select the desired format.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **TXT** tab.

The saving options on this tab are grouped into the following categories:

Retain layout

Depending on how you are planning to use your electronic document, select the best option below:

- **Formatted text**
Retains paragraphs, but does not retain spacing and the exact positioning of objects. Additionally, blank spaces are used to reproduce paragraph indentations and tables. The resulting text will be left-aligned (right-to-left texts will be right-aligned).

Note: Vertical texts will be changed to horizontal in this mode.

- **Plain text**
Unlike the **Formatted text** mode, this mode does not retain formatting.

Text settings

- **Keep line breaks**
Retains the original arrangement of lines.
- **Insert page break character (#12) as page break**
Saves the original page arrangement.

- **Use blank line as paragraph separator**
Separates the paragraphs with blank lines.
- **Keep headers and footers**
Preserves headers and footers in the output document.

Character encoding

ABBYY FineReader detects the code page automatically. To change the code page or its type, select the desired code page or code page type from the drop-down lists under **Character encoding**.

Saving in CSV

To save your text in CSV:

- Click the arrow next to the "Save" button on the main toolbar, or click **File>Save Document As>CSV Document**.

Tip. Additional saving options are available in the **Options** dialog box: select **Tools>Options...**, click the **Save** tab, and then click the **CSV** tab.

The saving options on this tab are grouped into the following categories:

Text settings

- **Ignore text outside tables**
Saves only the tables and ignores the rest.
- **Insert page break character (#12) as page separator**
Saves the original page arrangement.
- **Field separator**
Selects the character that will separate the data columns in the CSV file.

Character encoding

ABBYY FineReader detects the code page automatically. To change the code page or its type, select the desired code page or code page type from the drop-down lists under **Character encoding**.

Saving E-books

To save your text in FB2 or EPUB:

- On the main toolbar, select **Formatted Text** from the drop-down list.
- Select **File>Save Document As>FictionBook (FB2)** or **>Electronic Publication (EPUB)**. Alternatively, you can use the saving button on the main toolbar. Click the arrow next to the saving button and select a format. If you do not see the format you need on the list, select **Save to Other Formats...** and select the desired format in the **Save As** dialog box.

Tip. You can select additional saving options on the **FB2/EPUB** tab of the **Options** dialog box (**Tools>Options...>Save>FB2/EPUB**).

This tab offers the following options:

Document settings

Here you can specify detailed information about the document: its title, authors, and keywords. You can also enter a comment in the **Annotation** field.

Picture settings

If your text contains very many pictures, the resulting file may be fairly large. The quality of the pictures and the size of the resulting file can be tweaked using the options in the **Picture settings** drop-down list.

Tip.

- To specify your own picture settings, select **Custom....** In the **Custom Picture Settings** dialog box, make the necessary settings and click **OK**.
- If you do not wish to retain any pictures, clear the **Keep pictures** option.

Retain layout

Select a layout retention method depending on how you are intending to use the document:

- **Formatted text**
Preserves paragraphs, but does not preserve exact positions of objects and line spacing. You will get continuous left-aligned text (texts printed in right-to left scripts will be right-aligned).

Note: Any vertical text will be displayed horizontally if this option is selected.

When saving to EPUB, you can additionally save fonts and font sizes by selecting **Save fonts and font sizes**.

- **Plain text**
Does not preserve text formatting.

Use first page as cover

Selecting this option will place the first page of the document on the cover of the e-book.

Embed fonts

Selecting this option will embed the saved fonts into the e-book.

Saving in DjVu

To save your text in DjVu:

- Select **File>Save Document As>DjVu Document**. Alternatively, you can use the saving button on the main toolbar. Click the arrow next to the saving button and select a format. If you do not see the format you need on the list, select **Save to Other Formats...** and select the desired format in the **Save As** dialog box.

Tip. You can select additional saving options on the **DjVu** tab of the **Options** dialog box (**Tools>Options...>Save>DjVu**).

This tab offers the following options:

Save mode

Select a saving method depending on how you are intending to use the document:

- **Text under the page image**
Saves the image of the page and places the recognized text in a separate invisible layer

underneath the image. Thus, you get a searchable DjVu document that looks almost exactly like the original.

- **Page image only**

Saves the image of the page. The output document will look almost exactly like the original, but the text in the document will not be searchable.

Multiple layers

The DjVu format uses a special compression technique which separates a page image into layers and applies different compression methods to each of them. By default, ABBYY FineReader will automatically determine whether multi-layer compression must be used on a page. To enable or disable multi-layer compression for a document, select or clear the **Automatic** (default) option in the **Multiple layers** drop-down list.

Picture settings

If your text contains very many pictures, the resulting file may be fairly large. The quality of the pictures and the size of the resulting file can be tweaked using the options in the **Picture settings** drop-down list.

Tip. To specify your own picture settings, select **Custom....** In the **Custom Picture Settings** dialog box, make the necessary settings and click **OK**.

Saving to Microsoft SharePoint (only for ABBYY FineReader Corporate Edition)

ABBYY FineReader allows you to save recognized text to a network resource — a website, an online library, etc.

To save recognized text in Microsoft SharePoint:

1. On the File menu, click Save to Microsoft SharePoint...
2. In the dialog box that opens, specify the file name, the format, the network folder the file will be saved to, and the file options.
3. Click **Save**.

Important!

- Make sure you have Microsoft Office SharePoint Portal Server Client Components installed on your computer and the Web Client is running.
- When working with addresses like **http://Server/share**, it is impossible to replace a file already existing in the network.

Saving an Image of the Page

To save one or more images:

1. Select the pages in the **Pages** window.
2. On the **File** menu, click **Save Images...**
3. In the **Save Images As** dialogue box, select the disk and folder to save the file to.
4. Select a graphical format you want to save the image in.

Important! If you want to save multiple pages to one file, choose the TIF format and select **Save as multi-page image file**.

5. Specify the file name and click **Save**.

Tip. When saving, you can select the image format and compression method. Image compression lets you decrease file size. Methods used for image compression differ by compression rate and data losses. Two factors define the choice of compression method: the resulting image quality and file size.

ABBYY FineReader lets you use the following compression methods:

- **ZIP** — provides for no data losses and is used for images containing large sections of one color. For example, this method is used for screenshots and black-and-white images.
- **JPEG** — used for gray and color images, for example, photos. This method provides for a high-density compression but leads to data losses and poorer image quality (blurred outlines and poor color palette).
- **CCITT Group 4** — provides for no data losses and is used for black-and-white images created in graphical programs or scanned images. The **CCITT Group 4** compression is a widespread compression method applied to virtually all images.
- **Packbits** — provides for no data losses and is used for scanned black-and-white images.
- **LZW** — provides for no data losses and is used for graphics and gray images.

Advanced Features

Chapter Contents

- Working in Other Applications
- Using Area Templates
- Recognition with Training
- User Languages and Language Groups
- Group Work in a LAN
- ABBYY Hot Folder


Working in Other Applications

When you install ABBYY FineReader 11, you may choose to integrate the program with Microsoft Office applications and Windows Explorer. Integration means that:

- An **ABBYY FineReader 11** toolbar will appear in Microsoft Word, Microsoft Excel, and Microsoft Outlook.
- An **ABBYY FineReader 11** option will appear in the supported format shortcut menu in Windows Explorer.

You will be able to review and edit recognition results using your usual Microsoft Office tools. You will also be able to launch quick tasks and custom automated tasks and open image and PDF files directly from Windows Explorer.

To perform OCR on a document in a Microsoft Office application:

1. Click  button on the **ABBYY FineReader 11** bar.
2. In the dialog box, check the options and click **Start**.

ABBYY FineReader will be launched, and the recognized text will be opened in the current Microsoft Office application when the OCR process is completed.

To open an image or PDF file from Windows Explorer:

1. Select the Windows Explorer file.
2. Select ABBYY FineReader 11>Open in ABBYY FineReader from the file's shortcut menu.

Note: This command will appear only if the program supports the selected file's format.

ABBYY FineReader 11 will launch and the selected image will be added to a new ABBYY FineReader document. If ABBYY FineReader 11 is already running, the image will be added to the current ABBYY FineReader document.

If the ABBYY FineReader 11 panel doesn't appear on the Microsoft Office application toolbar:

- Select **ABBYY FineReader 11** from the application toolbar's shortcut menu.

If the **ABBYY FineReader 11** option is not available on the shortcut menu, it means that ABBYY FineReader 11 integration with **Microsoft Office** applications was disabled during custom installation.

To enable integration:

1. Go to Add or remove programs (Start>Settings>Control Panel).

Note: In Microsoft Windows Vista and Windows 7, this command is called **Programs and Features**.

2. In the list of installed programs, select **ABBYY FineReader 11** and click **Change**.
3. In the **Custom Installation** dialog box, select the appropriate components.
4. Follow the setup instructions.

Using Area Templates

If you are processing a large number of documents with an identical layout (e.g. forms or questionnaires), analyzing the layout of every page will be very time consuming. To save time, you can analyze only one document in a set of similar documents and save the detected areas as a template. Then you can use this template for the other documents in the set.

To create an area template:

1. Open an image and either let the program analyze the layout automatically or draw the areas manually.
2. On the **Areas** menu, click **Save Area Template...** In the saving dialog box, provide a name for your template and click **Save**.

Important! To be able to use an area template, you must scan all the documents in the set using the same resolution value.

Applying an area template:

1. In the **Pages** window, select the pages to which you wish to apply an area template.
2. On the **Areas** menu, click **Load Area Template...**
3. In the **Open Area Template** dialog box, select the desired area template (area template files have the *.blk extension).
4. In the same dialog box, next to **Apply to**, select **Selected pages** to apply the template to the selected pages.

Note: Select **All pages** to apply the template to all the pages in the current ABBYY FineReader document.

5. Click the **Open** button.

Recognition with Training

The **Read with training** mode is useful for:

- Texts set in decorative fonts
- Texts that contain unusual characters (e.g. mathematical symbols)

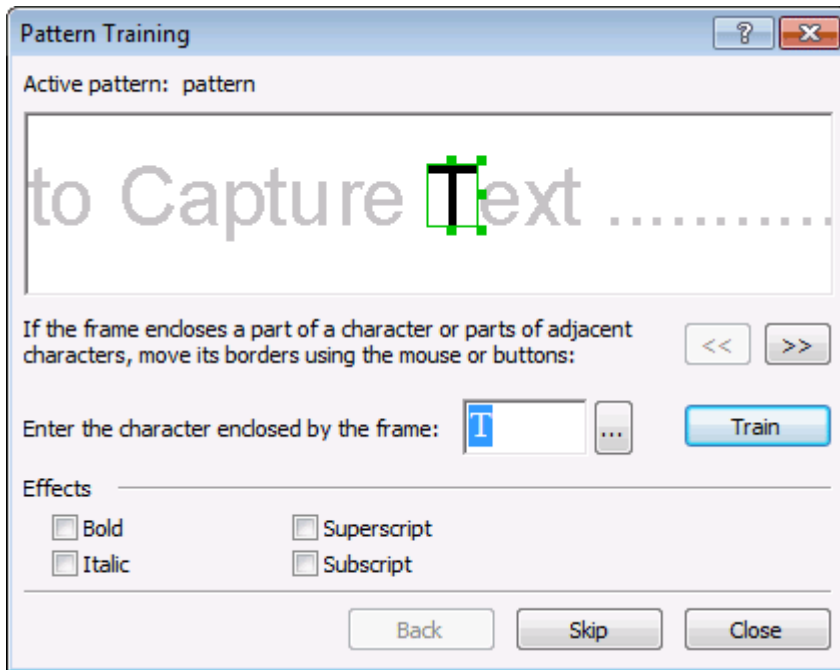
- Long (more than a hundred pages) documents of low print quality

Use the **Read with training** mode *only* if one of the above applies. In other instances, the slight increase in recognition quality will be outweighed by considerably longer processing times. For details, see Paper Document Contains Decorative (Non–Standard) Fonts.

- **Training User Patterns**
Provides instructions on training a user pattern, with additional information on training characters and ligatures.
- **Selecting a User Pattern**
Provides instructions on selecting a user pattern, with some additional tips on how to make the right choice.
- **Editing a User Pattern**
Provides instructions on viewing the characters in a user pattern and changing the font effects.

Training User Patterns

1. Select **Tools>Options...** and click the **Read** tab.
2. Under Training, select Use built-in and user patterns or Use only user patterns.
3. Select the **Read with training** option.
4. Click the **Pattern Editor...** button.
Note: Pattern training is not supported for Asian languages.
5. In the **Pattern Editor** dialog box, click **New...**
6. In the **Create Pattern** dialog box, type in a name for your new pattern and click **OK**.
7. Close the **Pattern Editor** and the **Options** dialog boxes by clicking the **OK** button in each.
8. On the toolbar at the top of the **Image** window, click **Read**.
Now if ABBYY FineReader encounters an unknown character, this character will be displayed in a **Pattern Training** dialog box.




9. Teach the program to read new **characters** and **ligatures**.

A ligature is a combination of two or three characters that are "glued together" (for example, fi, fl, ffi, etc.) and are difficult for the program to separate. In fact, better results can be obtained by treating them as single compound characters.

Note: Words printed in bold or italic type in your text or words in superscript/subscript may be retained in the recognized text by selecting the corresponding options under **Effects**.

To go back to a previously trained character, click the **Back** button. The frame will jump to its previous location and the latest trained "character image – keyboard character" correspondence will be removed from the pattern. The **Back** button navigates between characters of one word and will not navigate between words.

Important!

- You can only train ABBYY FineReader to read the characters included in the alphabet of the recognition language. To train the program to read characters that cannot be entered from the keyboard, use a combination of two characters to denote these non-existent characters or copy the desired character from the **Insert Character** dialog box (click  to open the dialog box).
- Each pattern may contain up to 1,000 new characters. However, avoid creating too many ligatures, as this may adversely affect OCR quality.

Selecting a User Pattern

ABBYY FineReader allows you to use patterns to improve OCR quality.

1. On the **Tools** menu, click **Pattern Editor....**
2. In the **Pattern Editor** dialog box, select the desired pattern from the list of available patterns and click **Set Active**.

Some important points to remember:

1. Rather than differentiating between some similar yet different characters, ABBYY FineReader recognizes them as one and the same character. For example, the straight ('), left ('), and right (') quotes will be stored in a pattern as a single character (straight quote). This means that left and right quotes will never be used in the recognized text, even if you try to train them.
2. For some character images, ABBYY FineReader will select the corresponding keyboard character based on the surrounding context. For example, an image of a small circle will be recognized as the letter O if there are letters immediately next to it and as the number 0 if there are digits next to it.
3. A pattern can only be used for documents that have the same font, font size, and resolution as the document used to create the pattern.
4. To be able to use a pattern later, save it to a file. See *Working with a FineReader Document* for details.
5. To recognize texts set in a different font, be sure to disable the user pattern by selecting the **Use only built-in patterns** option in **Tools>Options...>Read**.

Editing a User Pattern

You may wish to edit your newly created pattern before launching the OCR process. An incorrectly trained pattern may adversely affect OCR quality. A pattern should contain only whole characters or ligatures. Characters with cut edges and characters with incorrect letter correspondences should be removed from the pattern.

1. On the **Tools** menu, click **Pattern Editor....**
2. In the **Pattern Editor** dialog box, select the desired pattern and click the **Edit...** button.
3. In the **User Pattern** dialog box, select the desired character and click the **Properties...** button.
In the **Properties** dialog box, type the letter corresponding to the selected character and select the desired font effect (bold, italic, superscript, or subscript).

To delete a character that has been trained incorrectly, select the character and click the **Delete** button.


User Languages and Language Groups

You can create and use your own recognition languages and language groups in addition to the languages and groups included in the program.

- **Creating an OCR Language**
This is used to process texts that contain a lot of uncommon elements, such as code numbers.
- **Creating a Language Group**
This is used to recognize documents written in several languages that are not available from the predefined language combinations.

Creating an OCR Language

When performing OCR on a document, ABBYY FineReader uses some information about the language of the document (this language is to be selected from the **Document Languages** drop-down list in the main window). If there are too many unusual abbreviations or words in the text, the program may fail to recognize them correctly. If this is the case, you may wish to create your own recognition language for this document.

1. On the Tools menu, click Language Editor...
2. In the **Language Editor** dialog box, click **New...**
3. In the **New Language or Group** dialog box, select **Create a new language based on an existing one** and from the drop-down list below, select the desired language. Click **OK**.
4. In the **Language Properties** dialog box, specify the properties of the new OCR language.
 - a. **Language name** — Type a name for your OCR language in this field.
 - b. **Source language** — The language on which your new OCR language will be based. (Displays the language you selected in the **New Language or Group** dialog box. Click the arrow to the right to select a different language.)
 - c. **Alphabet** — Lists the characters of the alphabet of the source language. Click  to add or remove characters.
 - d. **Dictionary** — The dictionary that ABBYY FineReader will use to perform OCR on your document and to check the recognized text. The following options are available:
 - **None**
No dictionary will be used.
 - **Built-in dictionary**
The dictionary supplied with ABBYY FineReader will be used.
 - **User dictionary**
A user dictionary will be used. Click the **Edit...** button to add words to the dictionary or to import an existing user dictionary or text file in Windows (ANSI) or Unicode encoding. The words in the text file must be separated by spaces or other non-alphabetic characters.

Note: The words from the user dictionary may occur in texts in the following capitalizations: 1) lowercase only, 2) uppercase only, 3) first letter capitalized, 4) as spelt in the user dictionary. The four possibilities are summed up in the table below.

Word as spelt in the user dictionary	Possible forms of the word in texts
abc	abc, Abc, ABC
Abc	abc, Abc, ABC
ABC	abc, Abc, ABC
aBc	aBc, abc, Abc,

	ABC
--	-----

- **Regular expression**
You can use a regular expression to create a new language.
- e. **Advanced...** — Opens the **Advanced Language Properties** dialog box, where you can specify more advanced properties for your language:
 - Non-letter characters that may occur at the beginning or at the end of words
 - Standalone non-letter characters (punctuation marks, etc.)
 - Characters to be ignored if they occur inside words
 - Prohibited characters that may never occur in texts written in this language
 - All the characters of the language to be recognized
 - Allow Arabic numerals, Roman numerals, and abbreviations
- 5. Once you have finished creating your new language, select it as the recognition language for your document.

By default, user languages are saved in the ABBYY FineReader document folder. You can also save all your user languages and patterns to one file: Select **Tools>Options...**, click the **Read** tab, and then click the **Save to File...** button.

Creating a Language Group

If you are going to use a particular language combination regularly, you may wish to group the languages together for convenience.

1. On the **Tools** menu, click **Language Editor....**
2. In the **Language Editor** dialog box, click **New....**
3. In the New Language or Group dialog box, select Create a new group of languages and click OK.
4. In the **Language Group Properties** dialog box, type a name for your new group and select the desired languages.

Note: If you know that your text will not contain certain characters, you may wish to explicitly specify these so-called prohibited characters. Specifying prohibited characters can increase both recognition speed and quality. To specify prohibited characters, click the **Advanced...** button in the **Language Group Properties** dialog box. In the **Advanced Language Group Properties** dialog box, type the prohibited characters in the **Prohibited characters** field.

5. Click **OK**.

The newly created group will be added to the **Document Languages** drop-down list in the main window.

By default, user language groups are saved in the ABBYY FineReader document folder. You can also save all your user languages and patterns to one file: Select **Tools>Options...**, click the **Read** tab, and then click the **Save to File...** button.

Tip. If you need a particular language combination for a document, you can also select the desired languages directly, without creating a group.

1. From the **Document Languages** drop-down list, select **More languages...**
2. In the Language Editor dialog box, select Specify languages manually.
3. Select the desired languages and click **OK**.

Group Work in a LAN

ABBYY FineReader 11 provides tools and capabilities for processing documents over a local area network (LAN).

- Processing the Same ABBYY FineReader Document on Several LAN Computers
- Using the Same User Languages and Dictionaries on Several Computers

Processing the Same ABBYY FineReader Document on Several LAN Computers

When working with an ABBYY FineReader document over a network, you can use all the functionality available in ABBYY FineReader except the Style Editor.

If an ABBYY FineReader document is to be used on several LAN computers simultaneously, the following should be taken into account:

- A separate copy of ABBYY FineReader 11 should be installed on each computer.
- All the users must have full access to the ABBYY FineReader document.
- Each user may add pages to the document and modify them. If a user adds new pages and launches the OCR process for them, the program will process the entire document anew. None of the users will be able to edit the document while the processing is underway.

Once the page background changes to white, all the tools in the **Text** window will become available.

- A document page that is being processed by one user is locked for the other users.
- Users may monitor page status in the **Pages** window, including the computer that is processing the document, which user has opened a page, whether it has been recognized or edited, etc. To see the status, switch into the Details view.
- The fonts that are used in the ABBYY FineReader document must be installed on each computer. Otherwise, the recognized text may be displayed incorrectly.

Using the Same User Languages and Dictionaries on Several Computers

In ABBYY FineReader 11, several users can use the same user languages and dictionaries simultaneously. Several users can also add words to the same user dictionaries when working in the LAN.

To make user dictionaries and languages available to multiple users:

1. Create/open an ABBYY FineReader document and select desired scanning and OCR options for the document.
2. Specify a folder where user dictionaries will be stored. This folder should be accessible to all network users.

Note: By default, user dictionaries are stored in **%Userprofile%\Application Data\ABBYY\FineReader\11.00\UserDictionaries** if your computer is running Windows XP and in **%Userprofile%\AppData\Roaming\ABBYY\FineReader\11.00\UserDictionaries** if your computer is running Microsoft Windows Vista or Windows 7.

To specify the folder:

1. Select **Tools>Options...** and on the **Advanced** tab of the **Options** dialog box click the **User Dictionaries...** button.
2. Click **Browse...** and select the folder.

Save your user patterns and languages to an *.fbt file:

1. Select **Tools>Options...** and click the **Read** tab in the **Options** dialog box.
2. Under User patterns and languages, click the Save to File... button.
3. In the **Save Options** dialog box, specify a name for the *.fbt file and a folder where the file will be stored. This folder should be accessible to all network users.

Now when a user creates a new ABBYY FineReader document, they must load the user patterns and languages from the *.fbt file saved at step 3 and specify the path to the folder created at step 2, where the user dictionaries are stored.

Important! To be able to access the user dictionaries, patterns, and languages, users must have read/write permissions for the folders where they are stored.

To view the list of available user languages in the **Language Editor** dialog box (select **Tools>Language Editor...** and scroll to **User languages**).

When a user language is being used by several users, it is available as "read-only," and users will not be able to change user language properties. However, words can be added to and deleted from a user dictionary that is being used on several LAN computers.

When a dictionary is being edited by a user, it is available as "read-only" to other users, i.e. users can use this dictionary to perform OCR and check spelling, but they cannot add or remove words.

Any changes a user makes to a user dictionary will become available to all users who have selected the folder where this dictionary is stored. For the changes to come into effect, users will need to restart ABBYY FineReader.

ABBYY Hot Folder (ABBYY FineReader 11 Corporate Edition only)

ABBYY FineReader includes **ABBYY Hot Folder**, a scheduling agent which allows you to select a folder with images and set the time for ABBYY FineReader to process the images contained in the folder. For example, you can schedule your computer to recognize images overnight.

To process images in a folder automatically, create a processing task for that folder and specify the image opening, OCR, and saving options. You must also specify how often ABBYY FineReader should check the folder for new images (at regular intervals or only once) and set the start time for the task.

Important! At the time when the task should start, your computer must be turned on and you must be logged on.

- Installing and Starting ABBYY Hot Folder
- Main Window
- Creating a Task
- Changing Task Properties
- Results Log

Installing and Starting ABBYY Hot Folder

By default, ABBYY Hot Folder is installed on your computer together with ABBYY FineReader 11. If, during custom installation, you chose not to install ABBYY Hot Folder, you can install the application later by following the instructions below:

1. On the Windows taskbar, click **Start** and then select **Settings>Control Panel**.
2. Double-click the **Add or Remove Programs** icon in the **Control Panel** window.
3. In the list of installed programs, select **ABBYY FineReader 11** and then click **Change**.
4. In the **Custom Setup** dialog box, select **ABBYY Hot Folder**.
5. Follow the instructions of the setup program.

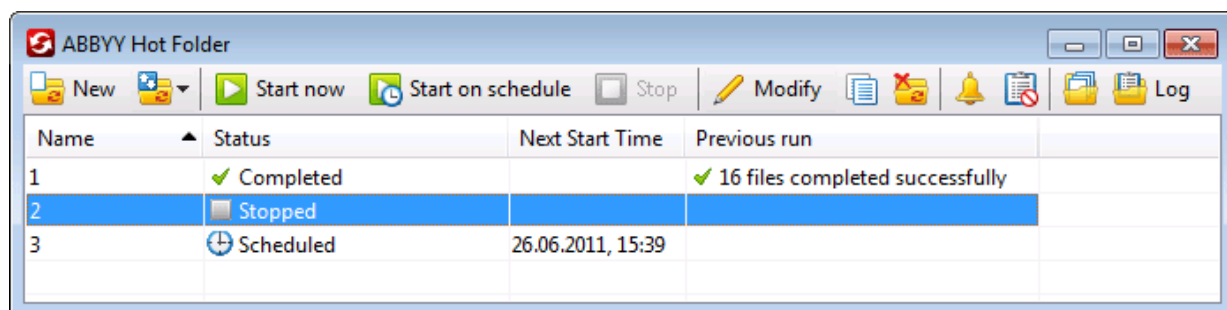
Starting ABBYY Hot Folder

There are several ways to start ABBYY Hot Folder:

- Select Start>All Programs>ABBYY FineReader 11 and then select ABBYY Hot Folder or
- In the ABBYY FineReader 11 main window, select Tools>Hot Folder... or
- Double-click the ABBYY Hot Folder icon on the Windows taskbar (this icon will appear once you have set up at least one hot folder).

Main Window


Starting ABBYY Hot Folder opens the feature's main window. Use the buttons on the toolbar to set up, copy, delete, and start folder processing tasks. You can also view detailed reports for each folder.



To create a new task, click the **New** button on the toolbar. The **ABBYY Hot Folder Wizard** will guide you through setting up a hot folder.






Note:

- By default, task files are stored in **%Userprofile%\Local Settings\Application Data\ABBYY\HotFolder\11.00**.
(In Microsoft Windows Vista, Windows 7 in **%Userprofile%\AppData\Local\ABBYY\HotFolder\11.00**).
- The results will be saved to the folder you specified when setting up the task. In some cases, to prevent the loss of already processed files, the program will create a subfolder in the specified folder and save the subsequent recognition results to this subfolder. The subfolder will be named as follows:
 1. **"Hot Folder DD.MM.YYYY HH-MM-SS"**
A subfolder with this name is created when you run a task again.
 2. **"File name(NUMBER)"**
A subfolder with this name is created if a file with a particular name already exists in the folder you specified.
- Depending on your version of the product, the number of pages that you can recognize within 30 days may be limited to 5,000.


To see messages notifying you that tasks have been completed, click  on the toolbar. A notification message will be displayed in a pop-up window above the Windows taskbar.

The ABBYY Hot Folder main window displays a list of set-up tasks. For each task, the full path to the corresponding hot folder is displayed, together with its current status and the scheduled processing time.

Tasks may have the following statuses:


Status	Description	
	Running...	The images in the folder are being processed.
	Scheduled	You chose to check the hot folder for images only once at start time. The start time is indicated in the Next Start Time column.
	Stopped	Processing has been stopped by the user.
	Completed	Finished processing images in this folder.
	Error (see log)	An error occurred when processing images in this folder. ABBYY FineReader did not complete the processing tasks. To find out the cause of the error, select the failed folder and click the View Log button on the toolbar.

You can save any task for future use:

1. Click the  button on the toolbar and select **Export...**

2. Specify the name of the task and the path to the task.
3. Click **Save**.

To load a previously saved task:

1. Click the  button on the toolbar and click **Import...**
2. Select the desired task and click **Open**.

Creating a Task

To create a task:

1. Either:
 - click the **New** button in ABBYY Hot Folder or
 - in the main ABBYY FineReader window, click the arrow next to the **Save** button and select the **Read Document Later...** command
A new task will be displayed, with the steps **Run once** and **Open FineReader document** already selected.
2. Select how often the task should be launched:
 - **Run once** (the task will be launched once at the time you specify)
 - **Recurring** (the task will be launched multiple times)
3. Specify the folder or document that contains the images to process.
 - **Open file from folder** allows you to process images in local and network folders.
 - **Open file from FTP** allows you to process images in folders on an FTP server. If login and password are required in order to access the FTP folder, enter them in the **FTP login** and **FTP password** fields respectively. If no login and password are required, select the **Anonymous connection** option.
 - **Open file from Outlook** allows you to process images in your mail box.
 - **Open FineReader document** opens images stored in a FineReader document.

Under **Documents**, select one of the three options:

- **Create a separate document for each file**
 - **Create a separate document for each folder**
 - **Create one document for all files**
4. Click **Analyze and read**.
 - **Analyze and read** analyzes an image and performs OCR.
 - **Analyze only** analyzes an image.
 - **Automatically analyze areas** detects text, picture, table, and other areas on the image.

- **Use template** loads a template with pre-configured areas.
- From the **Language** drop-down list, select the language of the texts to be recognized.

Note: You can select more than one language.

- Click the **All Document Settings...** button to select more options on the **Read** tab of the **Options** dialog box.

5. If required, select the **Save FineReader document** step.

6. Click **Save document** and specify the folder where to store the recognition results.

- From the **Save as** drop-down list, select a format for the output files.

Tip. You can have more than one saving step in order to save the results in multiple formats.

- In the **Folder** field, provide the path to the output folder where the recognition results will be stored.
- In the **File name** field, enter a name for the output files.

Note: The appearance of this dialog box will vary depending on the options selected at step 2.

7. Click **OK**.

The set-up task will be displayed in the ABBYY Hot Folder main window.

Note:

- By default, task files are stored in **%UserProfile%\Local Settings\Application Data\ABBYY\HotFolder\11.00**.
(In Microsoft Windows Vista, Windows 7 in **%UserProfile%\AppData\Local\ABBYY\HotFolder\11.00**).
- The results will be saved to the folder you specified when setting up the task. In some cases, to prevent the loss of already processed files, the program will create a subfolder in the specified folder and save the subsequent recognition results to this subfolder. The subfolder will be named as follows:
 1. **"Hot Folder DD.MM.YYYY HH-MM-SS"**
A subfolder with this name is created when you run a task again.
 2. **"File name(NUMBER)"**
A subfolder with this name is created if a file with a particular name already exists in the folder you specified.

Important! For tasks to start, your computer must be turned on and you must be logged on.

Changing Task Properties

You can change the properties of an existing task.

To change the properties of a task

1. Stop the task whose properties you wish to change.
2. On the ABBYY Hot Folder toolbar, click the **Modify** button.

3. Make the necessary changes and click **OK**.

Results Log

Images in a hot folder are processed according to your defined settings. ABBYY Hot Folder keeps a detailed log of all the operations.

The log file contains the following information:

- The name of the task and its settings
- Errors and warnings (if any)
- Statistics (number of processed pages, number of errors and warnings, number of uncertain characters)

To enable logging:

- On the ABBYY Hot Folder toolbar, click .

The log will be saved as a TXT file in the same folder that contains the files with recognized text.

To open a log file:

1. In the ABBYY Hot Folder main window, select the task you wish to inspect.
2. Click the **View Log** button on the toolbar.

Note: When you hover the mouse pointer over the status of the selected task, a log fragment is displayed in a pop-up window.

Appendix

Chapter Contents

- Glossary
- Supported Languages
- Supported Image Formats
- Supported Saving Formats
- Languages and Their Fonts
- Regular Expressions
- Shortcuts
- Using ABBYY FineReader Help

Glossary

A

ABBYY Business Card Reader is a handy application that enables users to scan business cards, capture the contact data, and export the captured contacts into various electronic formats. The application can also handle scans and photos of business cards stored on users' computers.

ABBYY FineReader document is an object created by ABBYY FineReader to process a paper document. It contains the images of the document pages, their recognized text (if any), and the program settings.

ABBYY Hot Folder is a scheduling agent which allows you to select a folder with images and specify the time for processing the images in this folder. At the specified time, the images from the selected folder will be processed automatically.

ABBYY Screenshot Reader is an application that enables you to create screenshots and recognize texts on them.

Abbreviation is a shortened form of a word or phrase used to represent the whole. For example, MS-DOS (for Microsoft Disk Operating System), UN (for United Nations), etc.

Activation is the process of obtaining a special code from ABBYY which allows the user to use his/her copy of the software in full mode on a given computer.

Activation code is a code that is issued by ABBYY to each user of ABBYY FineReader 11 during the activation procedure. An activation code is used to activate ABBYY FineReader on the computer that generated the **Product ID**.

Active area is a selected area on an image that can be deleted, moved or modified. To make an area active, click it. The frame enclosing an active area is bold and has small squares that can be dragged to change the size of the area.

Automatic Document Feeder (ADF) is a device that automatically feeds documents to a scanner. A scanner with an ADF can scan multiple pages without manual intervention. ABBYY FineReader supports multi-page documents.

ADRT® (Adaptive Document Recognition Technology) is a technology that increases the quality of conversion of multi-page documents. For example, it can recognize such structural elements as headings, headers and footers, footnotes, page numbering, and signatures.

Area is a section of an image enclosed by a frame and containing a certain type of data. Before performing OCR, ABBYY FineReader detects text, picture, table, and barcode areas in order to determine which sections of the image should be recognized and in what order.

Area template is a template that contains information about the size and location of the areas for a set of similar-looking documents.

B

Background picture area is an image area that contains a picture with text printed over it.

Barcode area is an image area that contains a barcode.

Brightness is a scanning parameter that indicates the contrast between the black and white areas on an image. Setting the correct brightness value increases recognition quality.

C

Code page is a table that establishes correspondences between characters and their codes. Users can select the characters they need from those available in a code page.

Color mode determines whether document colors are to be retained. Black-and-white images produce smaller FineReader documents and are faster to process.

Compound word is a word made up of two or more stems (general meaning). In ABBYY FineReader, a **compound word** is a word which is not in the dictionary but which the program thinks may be made up of two or more dictionary words.

D

Document analysis is a process of identifying the elements of the logical structure of a document and areas with different types of data. Document analysis can be carried out automatically or manually.

Document open password is a password which prevents users from opening a PDF document unless they type the password specified by the author.

Document options is the set of options that can be selected in the **Options** dialog box (**Tools>Options**). Document options also include user languages and patterns. Document options can be saved and then used in other ABBYY FineReader documents.

Dots per inch (dpi) is a measure of image resolution.

Driver is a software program that controls a computer peripheral (e.g., a scanner, a monitor, etc).

F

Font effects is the appearance of a font (i.e. bold, italic, underlined, strikethrough, subscript, superscript, small caps).

I

Ignored characters are any non-letter characters found in words (e.g. syllable characters or stress marks). These characters are ignored during the spell check.

Inverted image is an image with white characters against a dark background.

L

License Manager is a utility used for managing ABBYY FineReader licenses and activating ABBYY FineReader 11 Corporate Edition.

Ligature is a combination of two or more characters which are "glued together" (e.g. fi, fl, ffi). These characters are difficult for ABBYY FineReader to separate. Treating them as a single compound character improves OCR accuracy.

M

Monospaced font is a font (such as Courier New) in which all characters are equally spaced. For better OCR results on monospaced fonts, select **Tools>Options...**, click the **Document** tab, and select **Typewriter** under **Document print type**.

O

Omnifont system is a recognition system that recognizes characters set in any font and font size without prior training.

Optional hyphen is a hyphen (¬) that indicates exactly where a word or word combination should be split if it occurs at the end of a line (e.g. "autoformat" should be split into "auto-format"). ABBYY FineReader replaces all hyphens found in dictionary words with optional hyphens.

P

Page layout is the arrangement of text, tables, pictures, paragraphs, and columns on a page. The fonts, font sizes, font colors, text background, and text orientation are also part of the page layout.

Page layout analysis is the process of detecting areas on a page image. Areas can be of six types: text, picture, table, barcode, background picture, and recognition area. Page layout analysis can be performed automatically when clicking the **Read** button, or manually by the user prior to OCR.

Paradigm is the set of all grammatical forms of a word.

Pattern is a set of pairs of type "character image – actual character."

PDF security settings are restrictions that prevent a PDF document from being opened, edited, copied or printed. These settings include Document Open Passwords, Permissions Passwords, and encryption levels.

Permissions Password is a password which prevents other users from printing and editing a PDF document unless they type the password specified by the author. If some security settings are selected for the document, other users will not be able to change these settings until they type the password.

Picture area is an image area that contains a picture. This type of area may enclose an actual picture or any other object that should be displayed as a picture (e.g. a section of text).

Primary form is the "dictionary" form of a word (headwords of dictionary entries are usually given in their primary forms).

Print type is a parameter reflecting how the source text was printed (on a laser printer or a similar device, on a typewriter, etc.). For laser-printed texts, select **Auto**; for typewritten texts, select **Typewriter**; for faxes, select **Fax**.

Product ID is a parameter that is automatically generated on the basis of the hardware configuration when activating ABBYY FineReader on a given computer.

Prohibited characters — If certain characters will never occur in a text to be recognized, they may be included in a list of prohibited characters. Specifying prohibited characters increases the speed and quality of OCR.

R

Resolution is a scanning parameter measured in dots per inch (dpi). Resolution of 300 dpi should be used for texts set in 10 pt fonts and larger, 400 to 600 dpi is preferable for texts of smaller font sizes (9 pt and less).

Recognition area is an image area that ABBYY FineReader should analyze automatically.

S

Scanner is a device for inputting images into a computer.

Scanning mode is a scanning parameter that determines whether an image must be scanned in black and white, grayscale, or color.

Separators are symbols that can separate words (e.g. /, \, dash) and that are separated by spaces from the words themselves.

Support ID is a unique identifier of a serial number with information about the license and the computer on which it is used. A support ID provides additional protection and is checked by the technical support service before providing technical support.

T

Table area is an image area that contains data in table form. When the application reads this type of area, it draws vertical and horizontal separators inside the area to form a table. This area is then rendered as a table in the output text.

Tagged PDF is a PDF document which contains information about the document structure, such as its logical parts, pictures, and tables. The structure of a document is encoded in PDF tags. A PDF file with such tags may be reflowed to fit different screen sizes and will display well on handheld devices.

Task Manager is an ABBYY FineReader feature that allows you to run an automated task, create and modify automated tasks, and delete custom automated tasks which you no longer use.

Text area is an image area that contains text. Note that text areas should only contain single-column text.

Training is establishing a correspondence between a character image and the character itself. (For details, see the Recognition with Training section.)

U

Uncertain characters are characters that may have been recognized by the program incorrectly. ABBYY FineReader highlights uncertain characters.

Uncertain words are words containing one or more uncertain characters.

Unicode is a standard developed by the Unicode Consortium (Unicode, Inc.). The standard is a 16-bit international encoding system for processing texts. The standard is easily extended. The Unicode Standard determines the character encoding, the properties, and procedures used in processing texts written in a certain language.

Supported Image Formats

The table below lists the image formats supported in ABBYY FineReader 11.

Format	Extension	Open	Save
Bitmap	bmp, dib, rle	+	+
Bitmap, Black and White	bmp, dib, rle	+	+
Bitmap, Gray	bmp, dib, rle	+	+
Bitmap, Color	bmp, dib, rle	+	+
DCX	dcx	+	+
DCX, Black and White	dcx	+	+
DCX, Gray	dcx	+	+
DCX, Color	dcx	+	+
JPEG 2000	jp2, j2k	+	+
JPEG 2000, Gray	jp2, j2k	+	+
JPEG 2000, Color	jp2, j2k	+	+
JPEG	jpg, jpeg	+	+
JPEG, Gray	jpg, jpeg	+	+
JPEG, Color	jpg, jpeg	+	+

JBIG2	jb2, jbig2	+	+
PCX	pcx	+	+
PCX, Black and White	pcx	+	+
PCX, Gray	pcx	+	+
PCX, Color	pcx	+	+
PNG	png	+	+
PNG, Black and White	png	+	+
PNG, Gray	png	+	+
PNG, Color	png	+	+
TIFF	tif, tiff	+	+
TIFF, Black and White, Unpacked	tif, tiff	+	+
TIFF, Black and White, Packbits	tif, tiff	+	+
TIFF, Black and White, CCITT Group 4	tif, tiff	+	+
TIFF, Black and White, ZIP compression	tif, tiff	+	+
TIFF, Black and White, LZW compression	tif, tiff	+	+
TIFF, Gray, Unpacked	tif, tiff	+	+
TIFF, Gray, Packbits	tif, tiff	+	+

TIFF, Gray, JPEG compression	tif, tiff	+	+
TIFF, Gray, ZIP compression	tif, tiff	+	+
TIFF, Gray, LZW compression	tif, tiff	+	+
TIFF, Color, Unpacked	tif, tiff	+	+
TIFF, Color, Packbits	tif, tiff	+	+
TIFF, Color, JPEG compression	tif, tiff	+	+
TIFF, Color, ZIP compression	tif, tiff	+	+
TIFF, Color, LZW compression	tif, tiff	+	+
PDF	pdf	+	+
PDF v. 1.7 or earlier	pdf	+	+
DjVu	djvu, djv	+	+
GIF	gif	+	–
XPS (Microsoft .NET Framework 3.0 required)	xps	+	–
Windows Media Photo	wdp, wmp	+	–

Supported Saving Formats

ABBYY FineReader can save recognized texts in any of the following formats:

- Microsoft Word Document (*.doc)
- Microsoft Office Word 2007 Document (*.docx)
- Rich Text Format (*.rtf)

- OpenDocument Text (*.**odt**)
- Adobe Acrobat Document (*.**pdf**)
- HTML Document (*.**htm**)
- FB2 Document (*.**fb2**)
- EPUB Document (*.**epub**)
- Microsoft PowerPoint 2007 Presentation (*.**pptx**)
- Microsoft Office Excel Comma-Separated Values File (*.**csv**)
- Text Document (*.**txt**)
ABBYY FineReader supports Windows, DOS, Mac, and ISO codepages and Unicode encoding.
- Microsoft Excel Worksheet (*.**xls**)
- Microsoft Office Excel 2007 Workbook (*.**xlsx**)
- PDF/A (*.**pdf**)
- DjVu Document (*.**djvu**)

Languages and Their Fonts

Language	Font
Abkhaz	Arial Unicode MS ^(*)
Adyghian	Arial Unicode MS ^(*) , Lucida Sans Unicode
Agul	Arial Unicode MS ^(*) , Lucida Sans Unicode
Altai	Arial Unicode MS ^(*) , Lucida Sans Unicode
Arabic	Arial Unicode MS ^(*)
Armenian (Eastern, Western, Grabar)	Arial Unicode MS ^(*)
Avar	Arial Unicode MS ^(*) , Lucida Sans Unicode
Bashkir	Arial Unicode MS ^(*) , Palatino Linotype
Chechen	Arial Unicode MS ^(*) , Lucida Sans Unicode

Chinese (Simplified), Chinese (Traditional)	Arial Unicode MS ^(*) , SimSun fonts For example: SimSun (Founder Extended), SimSun-18030, NSimSun. Simhei, YouYuan, PMingLiU, MingLiU, Ming(for-ISO10646), STSong
Chuvash	Arial Unicode MS ^(*)
Chukchee	Arial Unicode MS ^(*) , Lucida Sans Unicode
Dargwa	Arial Unicode MS ^(*) , Lucida Sans Unicode
Ingush	Arial Unicode MS ^(*) , Lucida Sans Unicode
Gagauz	Arial Unicode MS ^(*)
Hausa	Arial Unicode MS ^(*) , Lucida Sans Unicode
Hebrew	Arial Unicode MS ^(*)
Japanese	Arial Unicode MS ^(*) , SimSun fonts For example: SimSun (Founder Extended), SimSun-18030, NSimSun. Simhei, YouYuan, PMingLiU, MingLiU, Ming(for-ISO10646), STSong
Kabardian	Arial Unicode MS ^(*) , Lucida Sans Unicode
Khakass	Arial Unicode MS ^(*)
Khanty	Arial Unicode MS ^(*)
Korean, Korean (Hangul)	Arial Unicode MS ^(*) , SimSun fonts For example: SimSun (Founder Extended), SimSun-18030, NSimSun. Simhei, YouYuan, PMingLiU, MingLiU, Ming(for-ISO10646), STSong
Koryak	Arial Unicode MS ^(*) , Lucida Sans Unicode

Lak	Arial Unicode MS ^(*) , Lucida Sans Unicode
Lezgi	Arial Unicode MS ^(*) , Lucida Sans Unicode
Mansi	Arial Unicode MS ^(*) , Lucida Sans Unicode
Mari	Arial Unicode MS ^(*)
Ossetian	Arial Unicode MS ^(*)
Russian (old spelling)	Arial Unicode MS ^(*) , Palatino Linotype
Tabasaran	Arial Unicode MS ^(*) , Lucida Sans Unicode
Tajik	Arial Unicode MS ^(*) , Palatino Linotype
Thai	Arial Unicode MS ^(*) , Aharoni, David, Levenim mt, Miriam, Narkisim, Rod
Udmurt	Arial Unicode MS ^(*)
Vietnamese	Arial Unicode MS ^(*)
Zulu	Arial Unicode MS ^(*) , Lucida Sans Unicode
Yakut	Arial Unicode MS ^(*)
Yiddish	Arial Unicode MS ^(*)

Where to find/supplied with

(*) Microsoft Office 2000 or later

Regular Expressions

The table below lists the regular expressions that can be used to create a new language.

Item name	Conventional regular expression symbol	Usage examples and explanations
-----------	--	---------------------------------

Any Character	.	c.t — denotes "cat," "cot," etc.
Character from Group	[]	[b-d]ell — denotes "bell," "cell," "dell," etc. [ty]ell — denotes "tell" and "yell"
Character not from Group	[^]	[^y]ell — denotes "dell," "cell," "tell," but forbids "yell" [^n-s]ell — denotes "bell," "cell," but forbids "nell," "oell," "pell," "qell," "rell," and "sell"
Or		c(a u)t — denotes "cat" and "cut"
0 or More Matches	*	10* — denotes numbers 1, 10, 100, 1000, etc.
1 or More Matches	+	10+ — allows numbers 10, 100, 1000, etc., but forbids 1
Letter or Digit	[0-9a-zA-Za-яA-Я]	[0-9a-zA-Za-яA-Я] — allows any single character [0-9a-zA-Za-яA-Я]+ — allows any word
Capital Latin Letter	[A-Z]	
Small Latin Letter	[a-z]	
Capital Cyrillic letter	[А-Я]	
Small Cyrillic letter	[а-я]	
Digit	[0-9]	
Space	\s	
	@	Reserved.

Note:

1. To use a regular expression symbol as a normal character, precede it with a backslash. For example, `[t-v]x+` stands for `tx`, `txx`, `txx`, etc., `ux`, `uxx`, etc., but `\[t-v]x+` stands for `[t-v]x`, `[t-v]xx`, `[t-v]xxx`, etc.
2. To group regular expression elements, use brackets. For example, `(a|b)+|c` stands for `c` or any combinations like `abbbbaabbb`, `ababab`, etc. (a word of any non-zero length in which there may be any number of a's and b's in any order), while `a|b+|c` stands for `a`, `c`, and `b`, `bb`, `bbb`, etc.

Examples

Suppose you are recognizing a table with three columns: birth dates, names, and e-mail addresses. In this case, you can create two new languages, Data and Address, and specify the following regular expressions for them.

Regular expression for dates:

The number denoting a day may consist of one digit (1, 2, etc.) or two digits (02, 12), but it cannot be zero (00 or 0). The regular expression for the day should then look like this: `((|0)[1-9])|([1|2][0-9])|(30)|(31)`.

The regular expression for the month should look like this: `((|0)[1-9])|(10)|(11)|(12)`.

The regular expression for the year should look like this: `([19][0-9][0-9]|([0-9][0-9])|([20][0-9][0-9]|([0-9][0-9])|([0-9][0-9])`.

Now all we need to do is combine all this together and separate the numbers by period (e.g. 1.03.1999). The period is a regular expression symbol, so you must put a backslash (`\`) before it. The regular expression for the full date should then look like this:

`((|0)[1-9])|([1|2][0-9])|(30)|(31)\.((|0)[1-9])|(10)|(11)|(12)\.((19)[0-9][0-9]|([0-9][0-9])|([20][0-9][0-9]|([0-9][0-9])|([0-9][0-9])`

Regular expression for e-mail addresses:

`[a-zA-Z0-9_-\.\+]\@[a-zA-Z0-9_-\.\+]`

Shortcut Keys

ABBYY FineReader 11 offers the following keyboard shortcuts for the most commonly used commands. You can also create your own keyboard shortcuts.

To create a shortcut:

1. Click **Tools>Customize....** The **Customize Toolbars and Shortcuts** dialog box will open.
2. On the **Keyboard Shortcuts** tab, in the **Categories** field, select a category.
3. In the **Commands** field, select the command for which you want to create keyboard shortcut or whose shortcut you want to change.
4. Click inside the **Press new shortcut key** field and then, on your keyboard, press the keys you want to use as a shortcut for the selected command.
5. Click **Assign**. The keys you specified will be added into the **Current shortcut** box.
6. Click **OK** to save the changes.

7. To return shortcuts to their default values, click **Reset** (for a selected command category) or **Reset All** (for all shortcuts).

The File Menu

Command	Shortcut
Scan Pages...	CTRL+K
Open PDF File/Image...	CTRL+O
New FineReader Document	CTRL+N
Open FineReader Document...	CTRL+SHIFT+N
Save Document As	CTRL+S
Save Images...	CTRL+ALT+S
Send FineReader Document by e-mail	CTRL+M
Send page images by e-mail	CTRL+ALT+M
Print Image	CTRL+ALT+P
Print Text	CTRL+P

The Edit Menu

Command	Shortcut
Undo	CTRL+Z
Redo	CTRL+ENTER
Cut	CTRL+X

Copy	CTRL+C CTRL+INSERT
Paste	CTRL+V SHIFT+INSERT
Clear	DELETE
Select All	CTRL+A
Find...	CTRL+F
Find Next	F3
Replace...	CTRL+H

The View Menu

Command	Shortcut
Show Pages Window	F5
Show Page Image Only	F6
Show Page Image and Page Text	F7
Show Page Text Only	F8
Show Zoom Window	CTRL+F5
Next Window	CTRL+TAB
Previous Window	CTRL+SHIFT+TAB
Properties...	ALT+ENTER

The Document Menu

Command	Shortcut
Read	CTRL+SHIFT+R
Analyze Layout	CTRL+SHIFT+E
Go to Next Page	ALT+DOWN ARROW PAGE DOWN
Go to Previous Page	ALT+UP ARROW PAGE UP
Go to Page...	CTRL+G
Close Current Page	CTRL+F4

The Page Menu

Command	Shortcut
Read Page	CTRL+R
Analyze Page Layout	CTRL+E
Edit Page Image...	CTRL+SHIFT+C
Delete All Areas and Text	CTRL+DELETE
Delete Text	CTRL+SHIFT+DELETE
Page Properties...	ALT+ENTER

The Areas Menu

Command	Shortcut
Read Area	CTRL+SHIFT+B
Change Area Type to Recognition Area	CTRL+1
Change Area Type to Text	CTRL+2
Change Area Type to Table	CTRL+3
Change Area Type to Picture	CTRL+4
Change Area Type to Barcode	CTRL+5
Change Area Type to Background Picture	CTRL+6

The Tools Menu

Command	Shortcut
Task Manager...	CTRL+T
Hot Folder...	CTRL+SHIFT+H
View Dictionaries...	CTRL+ALT+D
Language Editor...	CTRL+SHIFT+L
Pattern Editor...	CTRL+SHIFT+A
Verification...	CTRL+F7
Next Error	SHIFT+F4
Previous Error	SHIFT+F5

Options...	CTRL+SHIFT+O
------------	--------------

The Help Menu

Command	Shortcut
ABBYY FineReader 11 Help	F1

General

Command	Shortcut
Mark the selected text fragment bold	CTRL+B
Mark the selected text fragment <i>italic</i>	CTRL+I
<u>Underline</u> the selected text fragment	CTRL+U
Select table cells	Arrow keys
Select the Pages window	ALT+1
Select the Image window	ALT+2
Select the Text window	ALT+3
Select the Zoom window	ALT+4

How to Buy an ABBYY Product

You can buy ABBYY products from our online store or from our partners (see <http://www.abbyy.com> for the list of ABBYY partners).

For detailed information About ABBYY products, please

- visit our website at <http://www.abbyy.com>
- send us an e-mail to sales@abbyy.com
- contact the nearest ABBYY office

Additional fonts for various languages can be purchased from www.paratype.com/shop/.

About ABBYY

ABBYY is one of the world's leading developers that offers software and services for optical character recognition, document capture, linguistics, and translation.

ABBYY's products include:

For individual home and office users:

- ABBYY FineReader, an optical character recognition system
- ABBYY Lingvo, electronic dictionaries for desktop computers and mobile devices
- ABBYY PDF Transformer, a document conversion application that converts PDF files into editable formats and saves documents in PDF

For enterprises:

- ABBYY FlexiCapture, an all-in-one data capture solution
- ABBYY Recognition Server, a server solution that automates optical character recognition and PDF conversion processes
- ABBYY FineReader Bank, a data capture solution for processing payment orders and other banking documents

For developers:

- ABBYY FineReader Engine, a development kit for integrating OCR and document conversion technologies into Windows, Linux, MacOS, and FreeBSD applications
- ABBYY FlexiCapture Engine, a development kit for integrating data capture and document processing technologies into Windows applications
- ABBYY Mobile OCR SDK, a development kit for creating compact text recognition applications for various mobile platforms.

Throughout the world, organizations working with printed documents use ABBYY software to automate labor-intensive work and optimize business processes. ABBYY's products are used in large-scale government projects, for example by Australian Taxation Office, Tax Inspection of the

republic of Latvia, Ministry of Education of the Russian Federation, Ministry of Education of Ukraine and Government of Montgomery County, USA.

Technologies developed by ABBYY are licensed by the following companies: BancTec, Canon, EMC/Captiva, Hewlett-Packard, Microsoft, NewSoft, Notable Solutions, Samsung Electronics. Optical character recognition software is bundled with equipment of world's leading manufacturers: BenQ, Epson, Fujitsu, Fuji Xerox, Microtek, Panasonic, Plustek, Toshiba, Xerox and other.

The company's headquarters are located in Moscow, and regional offices are situated in Germany, USA, Japan, Ukraine and Taiwan.

For more information on ABBYY and its products, please visit www.abbyy.com.

Activating and Registering ABBYY FineReader

Software piracy hurts software manufacturers and end users alike – using an illegal product is never safe. Legal software ensures that a third party cannot introduce detrimental code changes. ABBYY makes every effort to protect its intellectual property rights and the security of its customers through a variety of anti-piracy measures.

ABBYY software products include special protection technology which prevents unauthorized use by those who have not entered into a license agreement with ABBYY.

To be able to run ABBYY FineReader in full mode, you need to activate it.

Registration is optional and provides a number of benefits.

Chapter Contents

- ABBYY FineReader Activation
- ABBYY FineReader Registration
- Privacy Policy

ABBYY FineReader Activation

After you install ABBYY FineReader 11, you need to activate the program to be able to run it in full mode. In restricted mode, some of the features will be disabled, depending on your version and region.

Activation takes just a few minutes. An Activation Wizard will help you send the data required to activate your copy to ABBYY, and you will receive an activation code in return.

Important! Some versions of the product are activated automatically over the Internet and do not require additional confirmation by the user.

To activate ABBYY FineReader:

1. On the Help menu, click Activate ABBYY FineReader....
2. Follow the instructions of the Activation Wizard.

When you choose to activate your copy by phone or e-mail, you will have to enter your activation code or provide the path to the activation file in the corresponding dialog box of the Activation Wizard.

Once you have activated your copy, ABBYY FineReader will run on this computer in full mode.

More about activation methods

- **Over the Internet**

Activation is carried out automatically and takes only seconds to complete. An active Internet connection is required for this method.

- **By e-mail**

The program will ask you to send an e-mail message to ABBYY with the information required to activate ABBYY FineReader. The e-mail message will be generated automatically. You will receive an activation code in reply.

Important! Be sure not to alter the subject field and the body of the generated e-mail message to ensure a quick reply from the e-mail robot.

- **By phone**

Select your country in the corresponding dialog box of the Activation Wizard. The wizard will show the telephone numbers of the nearest ABBYY office or ABBYY partner in your area. Communicate the serial number and the Product ID displayed in the Activation Wizard to an ABBYY technical support manager.

Important! Activation codes may be very long and may take some time to dictate over the phone.

- **Subscription**

Activation is carried out automatically if you choose to renew your license.

The Activation Wizard creates a unique Product ID. This parameter is created on the basis of your computer configuration data as of the moment of activation. The parameter does not include any personal information, information about the computer model, installed software, user settings, or stored data. In addition to the Product ID, the serial number, product name and version, and the interface language of ABBYY FineReader are transferred to the server during activation. This information is required in order to select the appropriate language and content of the activation reply message and will not be used for any other purposes.

You can re-install ABBYY FineReader on the computer on which it has been activated as often as you wish, without activating it again. The Activation Wizard detects and tolerates changes to your PC configuration. Minor upgrades will not require re-activation. If major upgrades are made to the system (e.g. if you format the hard disk or re-install the operating system), you may need to obtain a new activation code.

Once you activate ABBYY FineReader 11 Corporate Edition, you will also get access to ABBYY Business Card Reader, a handy application that enables you to scan business cards, capture the contact data, and export the captured contacts into various electronic formats.

ABBYY FineReader Registration

We recommend you to register your copy of ABBYY FineReader 11.

Registration benefits:

1. Free technical support
2. E-mail notifications about new ABBYY releases
3. BONUS APPLICATION! When you register ABBYY FineReader 11 Professional Edition, you get ABBYY Screenshot Reader for free. ABBYY Screenshot Reader creates screenshots of any part of your screen and performs OCR on areas that contain text.*

Note: Registration may be obligatory for some of the freeware versions of the product. If you have obtained a freeware version which requires registration, the product will not function unless you register your copy.

To register your copy of ABBYY FineReader, do one of the following:

Fill out the registration form when activating the program.

Note: If you do not register the program during activation, you can always do it later.

On the **Help** menu, click **Register Now...** and fill out the registration form.

Register your copy on the ABBYY website.

Contact the nearest ABBYY office.

* In ABBYY FineReader 11 Corporate Edition, ABBYY Screenshot Reader is already available and does not require registration.

Privacy Policy

You agree to transfer your personal data to ABBYY during registration of the ABBYY FineReader 11 software. You also agree that ABBYY may collect, process, and use the data received from you in accordance with terms of the License Agreement. Any personal information provided by you will be used only within the ABBYY group of companies and will not be disclosed to any third party unless ABBYY is required to so under applicable law or by the License Agreement.

ABBYY may send you e-mail messages containing product news, pricing information, special offers, and other company and product information only if you agree to receive such information from ABBYY by selecting the corresponding option in the product registration dialog box. You can delete your e-mail address from the mailing list at any time by contacting ABBYY.

Technical Support

If you have any questions regarding the use of ABBYY FineReader, please consult all the documentation you have (the User's Guide and Help) before contacting our technical support service. You may also wish to browse the technical support section on the ABBYY website at www.abbyy.com/support — you may find the answer to your question there.

Our technical support service experts will need the following information to answer your question competently:

- Your first and last name
- The name of your company or organization (if applicable)
- Your telephone number (or fax or e-mail)
- The serial number of your ABBYY FineReader package or support ID. Support ID is a unique identifier of the serial number, which contains information about your license and computer. To view your support ID, select **Help>About...>License Info**.
- Build number (see **Help>About...** and look up the **Build** field)
- A general description of your problem and the full text of the error message (if there was one)
- The type of your computer and processor
- The version of your Windows operating system
- Any other information you consider important

You can gather some of the above information automatically:

1. Select **Help>About...**
2. In the **About ABBYY FineReader** dialog box, click the **System Info...** button.

A dialog box will open displaying some of the above information.

Important! Only registered users are eligible for free technical support. You can register your copy of the application on the ABBYY website or by selecting **Help>Register Now...**