






ABBYY® FineReader® Engine 9.0

What's New

ABBYY FineReader Engine 9.0 is the next generation of ABBYY's award winning OCR technology. With Engine 9.0, ABBYY introduces **Adaptive Document Recognition Technology™ (ADRT™)**. ADRT revolutionises recognition technology by moving from a page-by-page to an entire document-based approach. ADRT analyses a document as a whole.

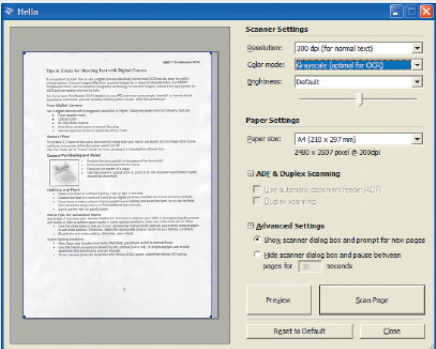
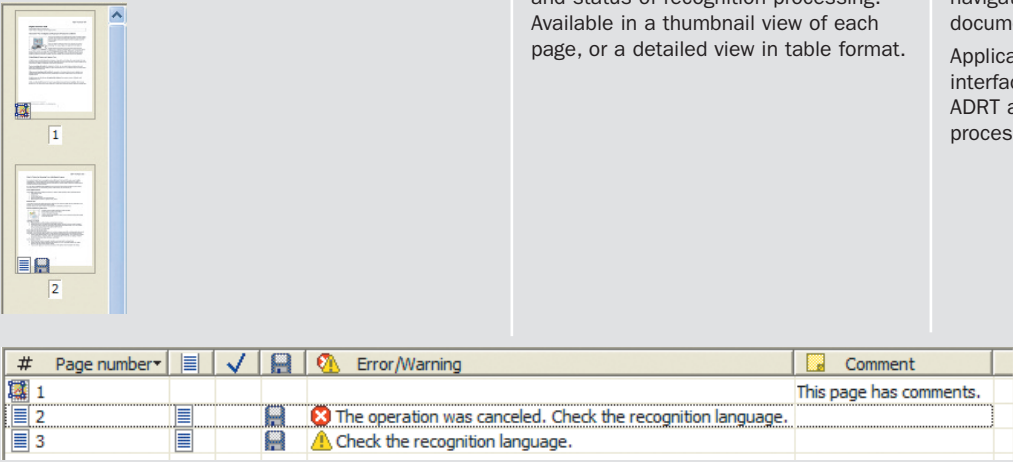
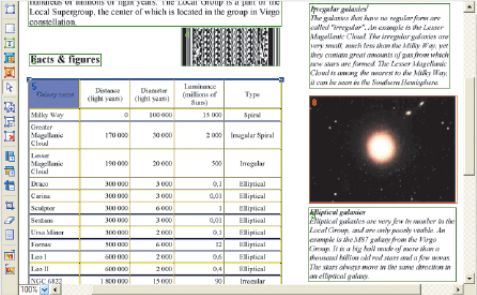
New features and enhancements for ABBYY FineReader Engine 9.0 include:

New Feature	Description	Benefit
Adaptive Document Recognition Technology (ADRT)	<p>New recognition technology based on an innovative set of document synthesis algorithms. Automatically builds a logical model of the document structure and identifies:</p> <ul style="list-style-type: none"> • Formatting purpose of elements like headers and footers, footnotes, page numbers etc. • How document elements should be reconstructed <p>ADRT automatically locates, determines and reconstructs various structural parts of the document and formatting of the elements.</p>	<p>ADRT is a major step forward in intelligent, easy to use document recognition.</p> <p>Documents generated by ADRT have consistent formatting across all pages of a document since they are processed as a unit.</p> <p>ADRT preserves integrity with logical relations between elements. For example, during the export to DOC(X) page numbers, headers and footers are automatically formatted and linked in an intelligent way.</p>
Multi-Page Processing through new Document specific API	<p>The new API objects allow you to set up the parameters of page and document synthesis separately. The pages will be processed as a logical unit and not isolated from each other. This approach preserves the logical organisation of the document, retaining not only the original text and columns, but also fonts, styles, etc.</p>	<p>Makes multi page document processing very simple, only 3 steps are needed: Open multiple pages or multi page documents, process and export them.</p>
Multi CPU / Multi Core Recognition Architecture	<p>Utilises all CPU cores during analysis and recognition of multi page documents.</p>	<p>Automatically combines and executes steps of distributing pages, as well as coordinating recognition and the synthesis. Provides easy scalability and utilisation of multi core hardware.</p>
New 2D barcode types   	<p>New support for additional barcode formats:</p> <p>Aztec 2D Barcodes* 2 dimensional matrix barcode which can contain up to 3000 characters</p> <p>Data Matrix 2D Barcodes* 2 dimensional matrix barcode capable of storing and encoding fifty characters in a symbol that is readable at size of 2 or 3 mm².The code can be read with only a 20% contrast ratio.</p> <p>QR Code 2D Barcodes* Next generation format used already in Japan for both image and text data with ability to store URLs which can be read and launched by appropriate software</p>	<p>The different barcode types are already established, for example:</p> <ul style="list-style-type: none"> • Aztec – often used on tickets by railways for example Deutsche Bahn and Swiss Federal Railways. • Data Matrix – labelling small items such as electronic components, also printed materials such as labels and letters. • QR Code – packaging, product and advertising info, and mobile-based applications. <p>The new 2D barcode support in FineReader Engine 9.0 opens new areas for your applications, like postal environments, ticket or voucher processing applications, document capturing with digital cameras.</p>
New Image pre-processing capabilities	<p>Additional image pre-processing functionalities:</p> <ul style="list-style-type: none"> • Detection of an image rotation up to 20 degrees • Deskew by horizontal and vertical pairs of black squares • Deskew by horizontal and vertical lines • Deskew by horizontal and vertical lines of text 	<p>Better OCR results on skewed images generated during scanning or document capture with a camera.</p>

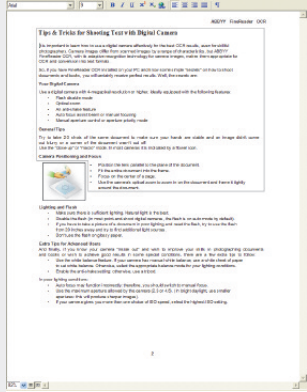
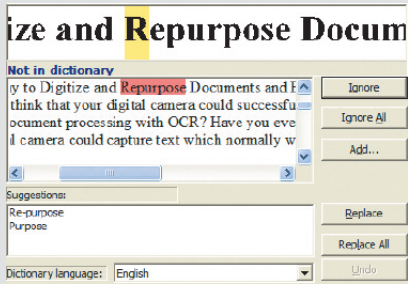
What's New

Visual Components

ABBYY FineReader Engine 9.0 Visual Components offer a set of graphical user interface components which developers can take advantage of when creating their own applications. FineReader Engine 9.0 Visual Components were created using ABBYY's expertise in end-user applications and are based on interfaces created for our award-winning FineReader Professional Edition application. The Visual Components are ActiveX controls for easy integration of user interface elements to existing applications. Developers can give users direct but controlled access to recognition results and functions for validation or checking of documents.

New Feature	Description	Benefit																																																																	
<p>Scan Interface</p> 	<p>Provides an interface for accessing TWAIN-based scanners. Allows users to select key scanning settings such as resolution, paper settings, etc.</p>	<p>Generic, consistent scan interface that works with different TWAIN compatible scanners.</p>																																																																	
<p>Document Viewer</p>  <table border="1" data-bbox="124 1417 1141 1529"> <thead> <tr> <th>#</th> <th>Page number</th> <th>Error/Warning</th> <th>Comment</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>1</td> <td></td> <td>This page has comments.</td> </tr> <tr> <td>2</td> <td>2</td> <td>The operation was canceled. Check the recognition language.</td> <td></td> </tr> <tr> <td>3</td> <td>3</td> <td>Check the recognition language.</td> <td></td> </tr> </tbody> </table>	#	Page number	Error/Warning	Comment	1	1		This page has comments.	2	2	The operation was canceled. Check the recognition language.		3	3	Check the recognition language.		<p>A visual overview of documents images and status of recognition processing. Available in a thumbnail view of each page, or a detailed view in table format.</p>	<p>Users can easily add, delete and navigate through the pages of the document.</p> <p>Applications can offer a great user interface to get the best results out of ADRT and document based processing.</p>																																																	
#	Page number	Error/Warning	Comment																																																																
1	1		This page has comments.																																																																
2	2	The operation was canceled. Check the recognition language.																																																																	
3	3	Check the recognition language.																																																																	
<p>Image Viewer</p>  <table border="1" data-bbox="124 1641 406 1921"> <thead> <tr> <th>Object name</th> <th>Distance (light years)</th> <th>Diameter (light years)</th> <th>Luminosity (millions of Suns)</th> <th>Type</th> </tr> </thead> <tbody> <tr> <td>Melita Wise</td> <td>0</td> <td>1000000</td> <td>15 000</td> <td>Spiral</td> </tr> <tr> <td>Centaur</td> <td>170000</td> <td>30000</td> <td>2000</td> <td>Irregular Spiral</td> </tr> <tr> <td>Cloud</td> <td>1500000</td> <td>20000</td> <td>500</td> <td>Irregular</td> </tr> <tr> <td>Dragon</td> <td>300000</td> <td>3000</td> <td>0.1</td> <td>Elliptical</td> </tr> <tr> <td>Carina</td> <td>300000</td> <td>3000</td> <td>0.01</td> <td>Elliptical</td> </tr> <tr> <td>Scorpion</td> <td>300000</td> <td>4000</td> <td>1</td> <td>Elliptical</td> </tr> <tr> <td>Sagittae</td> <td>300000</td> <td>3000</td> <td>0.01</td> <td>Elliptical</td> </tr> <tr> <td>Ursa Major</td> <td>300000</td> <td>2000</td> <td>0.1</td> <td>Elliptical</td> </tr> <tr> <td>Ursaes</td> <td>500000</td> <td>4000</td> <td>0.1</td> <td>Elliptical</td> </tr> <tr> <td>Leo I</td> <td>600000</td> <td>2000</td> <td>0.6</td> <td>Elliptical</td> </tr> <tr> <td>Leo II</td> <td>600000</td> <td>2000</td> <td>0.4</td> <td>Elliptical</td> </tr> <tr> <td>NGC 4032</td> <td>1 800 000</td> <td>15 000</td> <td>50</td> <td>Irregular</td> </tr> </tbody> </table>	Object name	Distance (light years)	Diameter (light years)	Luminosity (millions of Suns)	Type	Melita Wise	0	1000000	15 000	Spiral	Centaur	170000	30000	2000	Irregular Spiral	Cloud	1500000	20000	500	Irregular	Dragon	300000	3000	0.1	Elliptical	Carina	300000	3000	0.01	Elliptical	Scorpion	300000	4000	1	Elliptical	Sagittae	300000	3000	0.01	Elliptical	Ursa Major	300000	2000	0.1	Elliptical	Ursaes	500000	4000	0.1	Elliptical	Leo I	600000	2000	0.6	Elliptical	Leo II	600000	2000	0.4	Elliptical	NGC 4032	1 800 000	15 000	50	Irregular	<p>A full image view of a document page – it can be processed with specialised tools for creating, editing and changing the settings of recognition blocks. Tools include, for example:</p> <ul style="list-style-type: none"> Rotating, cropping and splitting images. Drawing recognition areas or selecting block types – text, image, table, barcode. Displaying or hiding toolbar buttons via code; customised buttons. 	<p>Gives users ability to influence:</p> <ul style="list-style-type: none"> Image view: zoom + rotate Image pre-processing: despeckle, invert Access to recognition results: block type, block order, read block Table processing results on rows and columns
Object name	Distance (light years)	Diameter (light years)	Luminosity (millions of Suns)	Type																																																															
Melita Wise	0	1000000	15 000	Spiral																																																															
Centaur	170000	30000	2000	Irregular Spiral																																																															
Cloud	1500000	20000	500	Irregular																																																															
Dragon	300000	3000	0.1	Elliptical																																																															
Carina	300000	3000	0.01	Elliptical																																																															
Scorpion	300000	4000	1	Elliptical																																																															
Sagittae	300000	3000	0.01	Elliptical																																																															
Ursa Major	300000	2000	0.1	Elliptical																																																															
Ursaes	500000	4000	0.1	Elliptical																																																															
Leo I	600000	2000	0.6	Elliptical																																																															
Leo II	600000	2000	0.4	Elliptical																																																															
NGC 4032	1 800 000	15 000	50	Irregular																																																															

Visual Components (continued)

New Feature	Description	Benefit
<p>Text Editor</p> 	<p>Provides option to highlight uncertain characters and basic text formatting tools. Developers can control the text are, available buttons, and other user actions in the text editor.</p>	<p>Offers a comfortable way for users to view and reformat recognised text prior to saving or export.</p>
<p>Text Validator</p> 	<p>An easy-to-use Wizard-like tool to check recognised characters marked as “uncertain.” Includes an integrated function to check spelling and provides a “zoom” view of key portions of text to be checked. Developers can control the behaviour of this component.</p>	<p>Offers an intuitive and comfortable way for users to check the spelling of the recognition results.</p>

Improved Asian Language OCR Support

New Feature	Description	Benefit
<p>ABBYY Chinese OCR</p> <p>ABBYY Japanese OCR</p> <p>ABBYY Korean OCR</p>	<p>New ABBYY technology for Chinese, Japanese and Korean languages. Functionality enhancements include enhanced accuracy, faster processing and the ability to combine these hieroglyphic Asian languages with other languages (until now only with English).</p>	<p>High accuracy and faster processing results for Asian documents. New possibility to work with multi-language documents like Chinese – French, Japanese – German etc.</p>

New Export Formats

New Feature	Description	Benefit
<p>XML-based Office 2007 File Formats</p>	<p>Export to new XML-based Microsoft Office formats – DOCX, XLSX and PPTX.</p>	<p>Users can export recognised documents to the new, open, interoperable, robust XML based formats that were introduced in Microsoft Office 2007.</p>
<p>MRC (Mixed Raster Content) Compression for PDF and PDF/A</p>	<p>Export to PDF with MRC compression. Supports the ability to set the MRC compression level. The parameters of compression for background, colour and text mask can be set and modified.</p>	<p>MRC compression achieves significantly better file compression without visible degradation of document representation. Significant reduced file size, up to 10 times smaller compared to JPEG compression. Ideal when colour documents are scanned and processed.</p>

What's New

New and Extended Licensing Options

New Feature	Description	Benefit
Extended CPU Core Support	Out-of-the box support for multi-core CPUs. New licensing scheme allows an unlimited number of cores with page limited licences.	Developers and customers can easily use multi-core hardware. Network licences allow flexible OCR distribution in an enterprise with increased speed and throughput.
CPU core based licences	New offer of licences without a page counter – instead pricing is based on the maximum number of CPU cores that can be used. Pricing is based on the maximum number of CPU cores that can be used instead.	CPU licences are often requested in server based processing scenarios.
Maximum Speed Limitation*	This new option sets maximum recognition speed limitation on characters per second.	If your business model is based on the maximum speed of the ORC/ICR Engine, then the new licensing model is perfect for your needs.

Architectural changes in FineReader Engine 9.0

With the release of our 9.0 platform products for developers, ABBYY will now offer separate development environments for document recognition/document conversion and data capture scenarios with ABBYY template matching technology.

ABBYY FineReader data capture functionality, which was previously offered as the FormReader batch and FlexiLayout processing add-ons for FineReader Engine 7.x/8.x, will soon be available through a separate Engine SDK – ABBYY FlexiCapture Engine. FlexiCapture Engine will support multi page fixed-form document templates and multi-page FlexiLayouts specifically suited for data capture environments. Please contact your ABBYY Sales representative for more information.

* Planned for Q4/2008 in maintenance release 1 of FineReader Engine 9.0

© 2008 ABBYY. All rights reserved. © 1987-2003 Adobe Systems Incorporated. All rights reserved. © 2001 ParaType Inc., Newton, Pragmatica, Courier fonts. Additional fonts for various languages can be bought on <http://www.paratype.com/shop/>. © 2003 ParaType Inc., Font OCR-B-GOST. Portions of this computer program are copyright © 1996-2007 LizardTech, Inc. All rights reserved. DjVu is protected by U.S. Patent No. 6,058,214. Foreign Patents Pending. © 1991-2007 Unicode, Inc. All rights reserved. This software is based in part on the work of the Independent JPEG Group.

ABBYY, the ABBYY logo, ABBYY FineReader, ADRT are either registered trademarks or trademarks of ABBYY Software Ltd. Adobe® PDF Library licensed from Adobe Systems Incorporated. Adobe, Acrobat, the Adobe Logo, the Acrobat Logo and the Adobe PDF logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries. Unicode is a trademark of Unicode, Inc. Microsoft, Excel, PowerPoint, Outlook, Windows are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.