



ABBYY® FineReader® Engine

Feature Overview Version 5.0-9.0

What's New in ABBYY FineReader Engine 5.0 (Released: 05/2001)

- Recognition quality improved by 1.5-2 times compared to 4.0 version
- Saves in HTML and PDF format with full page layout retention
- Full text color retention
- Recognition of subscript characters and simple chemical formulas
- Vertical text recognition and recognition of pictures "embedded" in table cells
- Dual page splitting
- 176 recognition languages, including programming languages Basic, C/C++, COBOL, Fortran, JAVA, Pascal, and new language dictionaries
- Component Object Model (COM) API accessible from any development environment supporting COM interface (Visual Basic, C/C++ etc.)
- API to create user languages and dictionaries
- Tools for training of user patterns for machine print characters via FineReader training dialog
- New HTML Help with context-sensitive topics accessible directly from VB Object Browser

What's New in ABBYY FineReader Engine 6.0 (Released: 08/2002)

- Improved algorithm for the recognition of poor print quality documents. The improved algorithm incorporates a new adaptive image binarization method and a new method of background removal, and is particularly effective in the case of images scanned in "grey" mode.
- New PDF saving mode - "Image only"
- Save text alignment in Excel format
- Save nonrectangular pictures in RTF format, recreate bullets and numbering
- 177 recognition languages
- New recognition fonts are supported: OCR-A, OCR-B and MICR (E13B).
- Fast mode available in all FineReader 6.0 Engine versions except the FineReader 6.0 Engine Standard. This mode provides faster recognition with worse image quality
- New features in ASCII version: the ability to preprocess image files, to recognize multipage image files, to work with memory images
- New Licence Manager utility

What's New in ABBYY FineReader Engine 7.0/7.1 (Released: 07/2004)

- Recognition quality improved by approximately 25%
- Opening and processing of PDF files
- New recognition languages: Traditional Chinese, Simplified Chinese and Japanese languages
- Old European languages have been added: Old English, Old French, Old German, Old Italian, and Old Spanish
- Recognition of Fraktur/Black Letter fonts
- Support for JPEG2000 part 1
- Opening a selected page from a multipage TIFF or PDF file
- New method for analysis and recognition of barcodes
- Support for new types of 1D barcodes: CODABAR without checksum, UCC Code 128, Industrial 2of5, IATA 2of5, Matrix 2of5, Code 93, UPC-A, and UPC-E
- Microsoft Word XML and ASCII XML output
- Export to MS PowerPoint
- Improved DA for invoices; detection of page orientation; 1D barcode detection, including detection of barcodes at any angle
- Improved detection and analysis of tables, particularly of tables without printed grid lines and tables with color rows and columns
- Improved adaptive binarization and background filtering
- New dictionaries added: law and medical dictionaries for the languages English and German
- Saving recognition results as linearized PDF files: the user will see the first pages of a PDF before the entire file has been downloaded
- Improved saving of edited texts in PDF format
- Numerous improvements of export to HTML and RTF formats
- Network runtime licences available
- Support for form and semi-structured document processing with support for ABBYY FormReader and FlexiCapture
- New recognition languages for ICR: Hungarian, Greek, and Croatian
- Arabic ICR digits
- Fast Mode for ICR



ABBYY® FineReader® Engine

Feature Overview Version 5.0-9.0

What's New in ABBYY FineReader Engine 8.0/8.1 (Released: 09/2005)

- Voting API support
- Field level recognition enhancements: fast mode for ICR, better text extraction from underlined fields, text block despeckling, better results on fields with spaces, dictionary with space-containing words
- PDF/A Support
- Up to 30 percent accuracy improvement on low resolution documents and faxes
- Up to 40 percent accuracy improvement on documents captured by using a digital camera
- Ability to straighten text lines on images taken by digital cameras
- New input image formats (GIF and DjVu)
- Balanced Processing Mode for OCR
- New Document Analysis for Full Text Indexing
- Improved PDF processing and creation, up to 2 times faster processing, accuracy improvement, enhanced security options, tagged PDF files, control of PDF page sizes
- Support for New Barcode Type – EAN 13 Supplemental
- CMC7 Text Type Support
- Additional Support for external dictionaries
- Forms and semi-structured documents processing improvements
- Ability to load Engine's subsystems on demand or preliminarily
- Ability to get all possible hypotheses for recognized words and characters
- Ability to trace Engine's calls in a log file
- "On the fly" core recognition tuning
- New Language for OCR: Thai
- New Language for OCR: Hebrew
- Expanded Asian Language Support for PDF and RTF Export
- Saving External Data in Engine Profiles

What's New in ABBYY FineReader Engine 9.0 (Released: 10/2008)

- **Adaptive Document Recognition Technology (ADRT):**
Documents generated by ADRT have consistent formatting across all pages of a document since they are processed as a unit.
- **Multi-Page processing through new Document specific API:**
The new API objects allow you to set up the parameters of page and document synthesis separately.
- **Multi CPU / Multi Core Recognition Architecture:**
Utilises all CPU cores during analysis and recognition of multi page documents.
- **New 2D barcode types: Aztec 2D*, Data Matrix 2D*, QR Code 2D***
- **New Image pre-processing capabilities:**
Detection of an image rotation up to 20 degrees, deskew by horizontal and vertical pairs of black squares, lines and lines of text
- **Visual Components - Scan Interface, Document Viewer, Image Viewer, Text Editor, Text Validator:**
Developers can give users direct but controlled access to recognition results and functions for validation or checking of documents.
- **Improved Asian Language OCR Support for Chinese, Japanese and Korean**
- **XML-based Office 2007 File Formats: DOCX, XLSX,**
Export recognised documents to the new, open, interoperable, robust XML based formats that were introduced in Microsoft Office 2007
- **MRC (Mixed Raster Content) Compression for PDF and PDF/A:**
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.
- **Licensing: Extended CPU Core Support**
New licensing scheme allows an unlimited number of cores with page limited licences.
- **Licensing: CPU core based licences**
New offer of licences without a page counter, Pricing is based on the maximum number of CPU cores that can be used instead
- **Licensing: Maximum Speed Limitation***
- **Data capture functionality,** which was previously offered as the FormReader batch and FlexiLayout processing add-ons will soon be available through a separate Engine SDK – ABBYY FlexiCapture Engine. Please contact your ABBYY Sales representative for more information.

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