ABBYY®

PRODUCT INFORMATION

ABBYY® FineReader Engine

Feature Overview Version 5.0-10

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
- API to create user languages and dictionaries
- 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.)
- 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 recognition of poor print quality documents. The improved algorithm incorporates a new adaptive image binarisation method and a new method of background removal, making it particularly effective with images scanned in greyscale.
- New PDF saving mode - "Image only"
- Saves text alignment in Excel format
- Saves nonrectangular pictures in RTF format, recreates bullets and numbering
- 177 recognition languages
- New recognition fonts: OCR-A, OCR-B and MICR (E13B).
- Fast mode available in all FineReader 6.0 Engine versions except the FineReader 6.0 Engine Standard.
- New features in ASCII version: the ability to preprocess image files, to recognise 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 recognition 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 binarisation and background filtering
- New dictionaries added: law and medical dictionaries for the languages English and German
- Saves recognition results as linearised 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 availability
- 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®

ABBYY® FineReader Engine

Feature Overview Version 5.0-10

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% accuracy improvement on low resolution documents and faxes
- Up to 40% 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
- Form and semi-structured documents processing improvements
- Ability to load Engine subsystems on demand or preliminarily
- Ability to get all possible hypotheses for recognised words and characters
- Ability to trace Engine 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
- Saves 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 Core 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.
- **Proprietary ABBYY 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 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 are now available through a separate Engine SDK – ABBYY FlexiCapture Engine. Please contact your ABBYY Sales representative for more information.

ABBYY®

ABBYY® FineReader Engine

Feature Overview Version 5.0-10

What's New in ABBYY FineReader Engine 10 (Released: 09/2010)

- **New Binarisation**
  The new binarisation ensures the whole text retention and prevents information losses even in difficult cases.
- **New Camera images pre-processing**
  With the new Camera OCR features your solution is ready to get the best OCR results out of paper documents captured with a camera.
- **Colour mask, mark and stamp filtering**
  An excellent feature for data capture systems, helps prevent loss of data from fields covered by stamps and colour marks.
- **Chart and diagram detection**
  Diagrams are better detected as a unit and it is possible to choose if the text should be extracted or if it should stay in the original image.
- **Speed: Tuned Fast Mode**
  New Enhanced Fast Mode designed to optimize processing speed/accuracy balance for images of good quality. Up to 92% faster page throughput while maintaining the high level of recognition accuracy**.
- **New mode for low resolution scans**
  Better recognition of critical low quality documents like faxes, 20% higher accuracy compared to the standard Normal Mode
- **Improved classifier for CJK**
  The recognition accuracy for Chinese, Japanese and Korean languages increased by up to 40% due to improved Asian characters classifier.
- **New Open Office Text Export***
  Support for ODT, an ISO standard and used by a lot of public institutions.
- **New E-book Formats***
  Two new e-book export formats: EPUB & FB2.
- **High quality, highly compressed PDF files & PDF Export Profiles**
  Predefined profiles with settings for frequently used export scenarios. For simplified fine tuning there are six new high level parameters, and 40 low-level ones.
- **Tuned profiles for popular usage scenarios**
  New optimised profiles for the most often executed recognition tasks.
- **Document structure API with Table of Contents reconstruction based on ADRT®**
  The new API gives developers access to and detailed information about the structure of a multi-page document. The table of content will also be exported automatically when the DOC(X) format is selected.
- **Improved Developer's Guide (Help)**
  Restructured documentation / help file. It now contains new content including a general product description, API specifications, usage samples and best practices.
- **New adjusted Project Pricing including Fraktur OCR**
  Revised project based pricing, perfectly suited for medium to large volume.
- **New CPU Core Pricing**
  The entry level of CPU Core Pricing was reduced.
- **New CPU Core Network Licensing**
  The CPU cores which execute OCR can now be distributed within a network, ideally suited for virtualised, server based environments.
- **More Default Runtime Features**
  New Add-Ons enable the addition of more features to your application without increasing the deployment costs, e.g. 2 D Barcode Support & Visual components.

* Still under development, planned to be released in a maintenance release of FineReader Engine 10 (Q1/2011)

** Based on internal testing of FineReader Engine 9 (release 1) vs. FineReader Engine 10 (release 1). Tests included standard business document scans in English, German, French, Spanish and Italian. Your results may vary based on scan quality, document complexity, system and application type.

ABBYY Europe GmbH
Elsenheimerstr. 49, 80687 Munich, Germany
Tel: +49 89 511159 - 0, Fax: +49 89 511159 - 59
sales_eu@abbyy.com
www.ABBYY.com