

ABBYY® Recognition Server 3.0

Professional Edition & Extended Edition



Robust OCR and Document Capture Solution

ABBYY Recognition Server, a robust, server-based solution, allows organisations to efficiently convert paper documents or document images into searchable and re-usable electronic files. Powerful technology supports document scanning, highly accurate full-text OCR, and metadata extraction in order to deliver content suitable for e-archives, enterprise content management, enterprise search and other document-based business processes. A scalable architecture enables high volume processing which can be deployed efficiently and managed centrally. Recognition Server's document capture and automated OCR functions can be easily installed and integrated with back-end systems and other applications.

Where to Use ABBYY Recognition Server?

Today, business success relies heavily on increasing productivity and extracting optimal use out of the information flooding the organisation each day. Leveraging cutting-edge technology, Recognition Server automates document handling processes and converts both paper and digital information into more usable formats, including PDFs with associated indexed data for easy retrieval and further use within business processes.

• Document Archiving and Back-log Conversion

Convert a large number of documents into searchable full-text formats for document management and archiving. Recognition Server 3.0 offers both full-text conversion and simple indexing of image-based documents. After converting documents to PDF, XML* and other formats, operators have the option to check recognition quality or index fields manually. The solution's easy-to-use integration tools help deliver data directly into an archive, SharePoint Server or other back-end systems.

• Everyday Conversion

As a server-based system, Recognition Server delivers around-the-clock access to OCR conversion without installing software on individual workstations. Users can simply send an e-mail or drag and drop documents into a specified folder to convert and deliver accurate, searchable results quickly. Recognition Server can also be linked to MFPs and network scanners to automatically process documents which are scanned or copied each day.

• Increased Productivity for Legal

Recognition Server 3.0 provides flexible and accurate technology for managing and cataloguing the influx of information coming into law offices. While much of electronic document evidence arrives in a searchable electronic format, email attachments of scanned PDF, JPEG or TIFF image files are not usually searchable – and therefore can slow down attempts to locate key pieces of data. Recognition Server quickly and easily converts these “locked” documents into searchable files (complete with Bates stamping and retained the native formatting).

• Enterprise Search Systems

Recognition Server 3.0 with the Google Search Appliance™ (GSA) and Microsoft® IFilter modules** works as a background OCR service to help organisations unlock and access documents saved in image-based formats. The solution gives enterprise organisations an automatic and cost-effective way to organise and access knowledge that was previously hidden from search results.

OCR and PDF Conversion via a Server-Based Architecture

- **Automated and Independent.** Schedule OCR to occur as documents arrive in “watched folders” or process documents at a scheduled time. Works as a background process separate from client workstations.
- **Multiple Points of Access.** Access documents from MFPs, workgroup scanners, the network, FTP-folders and e-mail inboxes.
- **High Volume Processing.** Parallel processing and automatic workload distribution maximises capacity.
- **Scalability.** Increase processing by adding workstations and cores.
- **Centralised Management.** Manage all set-up and administration centrally.
- **Accuracy.** More than 20 years of recognition technology experience delivers precisely accurate text conversion.
- **International.** Recognition in more than 190 languages.
- **Fault Tolerance.** Event logging, auto-start and re-start of OCR and re-routing of documents when required.
- **Flexible Integration.** XML-ticket Support, COM-based API and Web Service API enable integration with other applications.

* Specialised features available with Recognition Server Extended Edition only. Certain functions are available as Add-on modules for Extended Edition only.

** Available as Add-on Modules for Recognition Server Professional Edition, included in Recognition Server Extended Edition

Functionality Overview



Overview: Document Processing Steps

1. Scanning and Document Import:

- **Scanning:** Images are scanned by an operator on the Scanning Station and then sent to Recognition Server.
- **Document Import:** the Server Manager imports files from the designated input source (i.e. shared folder, FTP folder, or mailbox folder) and places them in a processing queue.

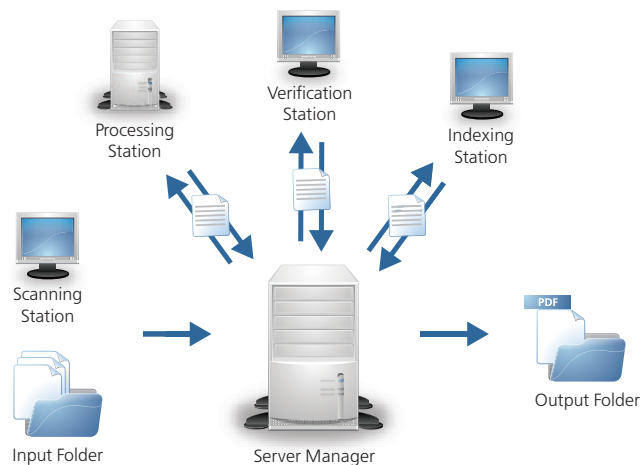
2. Recognition: The system evenly distributes files among the available Processing Stations for recognition. Workflow processing on all available Processing Stations is administered and monitored by the Server Manager. Both components function as Windows® services, can be installed separately in the network or on the same computer.

3. Verification (optional): In case of low quality document input or when precise accuracy is required, an optional Verification Station can be set up for manual proofing by an operator. The Server Manager routes documents to the Verification Stations according to workflow settings****.

4. Document Separation: After recognition and verification, documents can be sorted using blank page, barcode, or fixed page number separators. Scripts may also be used to create custom separation rules.

5. Classification and Indexing (optional) Organisations may choose to index manually on the Indexing Station, or index automatically using customised scripts which enable intelligent export and delivery using document properties and metadata.

6. Export: Following recognition and verification, the Server Manager delivers the output document to its destination, which can be a network folder, a SharePoint® library*, an e-mail address or applications which use the Recognition Server API.



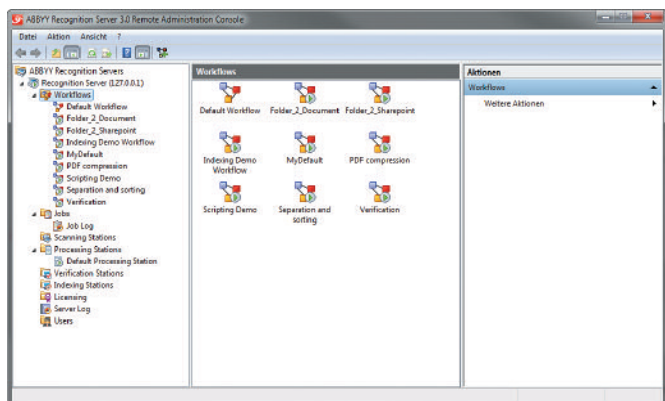
Set-up and Management

Management and Administration

ABBYY Recognition Server is administered remotely via the Microsoft Management Console (MMC). All system settings, including workflows, job lists, Processing Station properties, licences and server log files can be edited at a central location.

Conversion Workflows and Job Distribution

A Recognition Server workflow is a set of processing parameters which include document input source, processing stages and delivery methods. A job is the smallest processing unit and may include one or several images or PDF files. The administrator sets priorities for each workflow in advance to ensure that all jobs are processed in order of importance.



Scheduled and Efficient Processing

Administrators have the flexibility to set up particular workflows at specific times or intervals (i.e. daily, weekly, or monthly) in order to increase efficiency or take advantage of low-workload periods (i.e. night time). Administrators can also configure the system to specify that different Processing Stations be used at different times.

Scalability

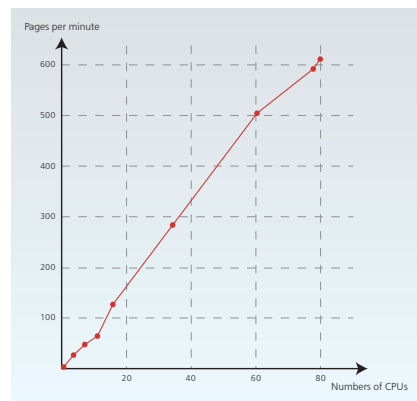
ABBYY Recognition Server can satisfy the needs of organisations of different sizes and document processing requirements. The Server Manager and Processing Stations can be installed on a multi-core system or distributed in the network. Either way, the flexible architecture scales efficiently to allow administrators to easily set up systems which can process as much as hundreds of pages per minute.

* Specialised features available with Recognition Server Extended Edition only. Certain functions are available as Add-on modules for Extended Edition only.

** Available as Add-on Modules for either Edition.

*** The diagram and processing capacity information is based on internal testing. System performance can vary depending on the quality of images, hardware performance, network configuration and other factors.

**** The Verification Station is scheduled for availability in Q1/2011.





Functionality Overview

Features

Input and Scanning

ABBYY Recognition Server offers an application for direct scanning to Recognition Server. The Scanning Station supports TWAIN, WIA and ISIS scanning interfaces and provides tools for visual quality control of documents before recognition. Pre-recognition tools available from the Scanning Station include document preview, image enhancement, batch type definition, manual redaction, and page order sorting.

Recognition and PDF Conversion

ABBYY's award-winning recognition technology, known for its accuracy and stability, underpins Recognition Server. The system supports a broad range of recognition functions and file input formats, including:

- **Image pre-processing:** split dual pages (for book scans), convert colour and grey to black and white, clear background noise, and more
- **Print Type:** Choose between normal text, typewriter, dot-matrix, OCR-A, OCR-B, and MICR (E13b)
- **Languages:** Select from more than 190 languages, accurately process even multilingual documents*
- **Special Fonts:** Support for black letter, Schwabacher and most other Gothic fonts printed between 1700 and 1937 in English, German, French, Italian and Spanish*
- **Barcodes:** Recognises most popular 1D and 2D barcodes including 2D Aztec, Data Matrix, and QR Code barcodes, no matter the angle or position on a document
- **3 different Processing Speed Modes:** precision, speed, and balanced mode

Customisation via Scripting

By leveraging scripting, users can customise internal processing, exceptions handling and document export. Different kinds of scripts can be used to create more flexible document separation rules, automate simple indexing, add dynamic and custom export functions and simplify publication of the output documents to external systems. Scripts can also be used for exception handling to manage processing of documents requiring special attention.

Quality Control and Verification****

Verification Stations, which can be installed on multiple workstations and controlled using a concurrent licence, enable operators to manually check or verify processing results. Verifiers can check the layout reconstruction (image, text, tables), edit the recognised text and check spelling.

Document Separation and Naming

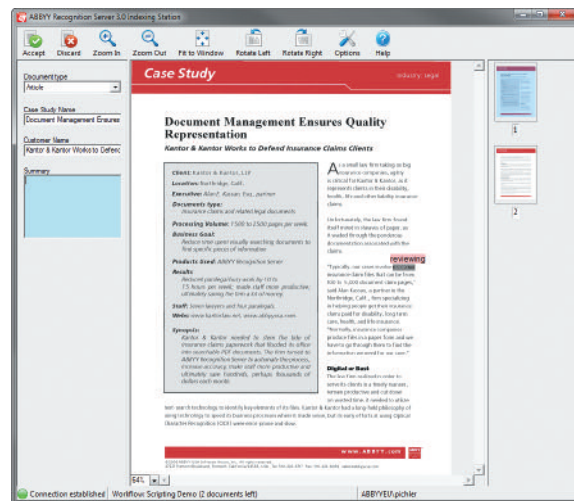
Recognition Server can use blank pages or barcodes as document separators. It can also create documents of a fixed number of pages. Documents in a sub-folder can also be merged into a single document. Users may also choose to create automatic naming conventions for documents using characteristics such as a date, time, specific text and barcode values. Using scripting controls, administrators can create custom rules for document separation.

Classification and Point-and-Click Indexing

The Indexing Station provides an interface which allows operators to specify document types and attributes. Operators can define indices in three different ways.

- **Type-in Indexing:** operators type in indices manually
- **“Point-and-click” Indexing:** operators extract keywords for indices by pointing and clicking on a document image
- **Key Phrase Indexing:** operators use a “rubber band” function for outlining phrases, multiple lines or paragraphs.

Administrators can also create and define automatic indexing functions based on document structure and content using scripts.



Output Formats

Recognition Server 3.0 outputs to a variety of formats, including DOC, DOCX, RTF, XML*, XLS, XLSX, HTML, TIFF, JPEG, JPEG2000 and more. It is possible to generate multiple output formats for one input document. Recognition Server supports export to “simple” searchable PDFs, linear and tagged PDFs, PDF files with security options and encryption and PDF/A for long-term archiving. It is also possible to generate highly compressed MRC-PDFs, which can effectively manage colour documents.

Export

Recognition and conversion results can be sent by e-mail or deposited in different locations such as ECM systems, archives or Sharepoint Servers.*

* Specialised features available with Recognition Server Extended Edition only. Certain functions are available as Add-on modules for Extended Edition only.

** Available as Add-on Modules for either Edition.

*** The diagram and processing capacity information is based on internal testing. System performance can vary depending on the quality of images, hardware performance, network configuration and other factors.

**** The Verification Station is scheduled for availability in Q1/2011.

Integration and Availability

General System Requirements:

- **PC** with Intel® Core™2/2 Quad/Pentium® /Celeron® /Xeon™, AMD K6/ Turion™/ Athlon™/ Duron™/ Sempron™ or compatible processor with a minimum clock speed of 2 GHz
- **Operating system:** Microsoft® Windows® 7, Windows Server® 2008, Windows Vista®, Windows Server 2003, Windows XP
- **Memory:** Server Manager; 1 GB RAM; Scanning Station; 1 GB RAM; Processing Station: 512 MB RAM, plus additional 300 MB for each recognition process; Indexing Station; 768 MB RAM
- **Hard Disk Space:** Server Manager: 20 MB for installation, additional 1 GB for program operation; Scanning Station: 1 GB; Processing Station: 600 MB for installation, additional 1 GB for program operation; Indexing Station: 500 MB for installation and 1 GB for program operation
- **Microsoft .NET Framework 2.0** or later for saving files to Microsoft SharePoint Server
- **Microsoft Outlook 2000** or later for processing mail messages via Microsoft Exchange Server
- **Microsoft IIS 5.1** or later (for Web API)
- **Scanner** supporting TWAIN, WIA or ISIS
- **Video card and display** with a resolution of 1024 x 768

Hard disk space required for program operation may be larger depending on complexity, quality, and number of images. System requirements may vary based on server component or additional module used. Contact your ABBYY sales representative for more detailed specifications.

Input Formats:

- BMP
- PCX
- JPEG
- JPEG 2000
- PNG
- TIFF
- PDF (up to PDF 1.6)
- DjVu

OCR Languages:

More than 190 languages including Chinese, Japanese, Korean, Hebrew

Output Formats:

- DOC (Microsoft Word)
- DOCX (Microsoft Word 2007)
- RTF
- PDF, PDF/A, PDF (Version 1.6)
- HTML
- CSV (comma separated values)
- TXT
- XLS (Microsoft Excel)
- XLSX (Microsoft Excel 2007)
- Original image files
- TIFF
- JPEG
- JPEG2000

Integration

ABBYY Recognition Server can easily be connected to external applications including Enterprise Content Management and Workflow systems, in order to extend existing applications and automate document conversion and delivery. Companies looking to implement Web services or other service oriented applications can also easily integrate Recognition Server.

Recognition Server offers a number of advanced integration options. Functionalities include:

- XML-Ticket Support*
- COM-based API*
- Web Service API*
- Microsoft SharePoint Export*

Connectors for Enterprise Search

ABBYY Recognition Server 3.0 expands the capabilities of Google and Microsoft search systems by creating a connection to the Google Search Appliance, Microsoft Office SharePoint Server, and Windows Desktop Search. These Add-on modules enable the search system to index the full text content of faxes, scanned agreements, correspondence, and any other TIFF, JPEG or PDF documents**.

Licensing

Recognition Server is sold in two product editions:

Recognition Server 3.0 Professional Edition

The Professional Edition offers standard functionalities for organisations which need an automated OCR background conversion service with partial adjustment and almost no integration with other applications.

Recognition Server 3.0 Extended Edition

Recognition Server 3.0 Extended Edition offers complete integration with external applications and implementation as part of a Web service architecture. Additional functionality offered in the Extended Edition include: additional OCR languages (Thai and Hebrew), XML-Export, Web Service API, COM-based API as well as support of XML-Tickets and Microsoft SharePoint Export. FineReader XIX (recognition of Black letter script) and OCR for Chinese, Japanese and Korean are available only for Recognition Server 3.0 Extended Edition.

Add-on Modules

Both editions can be extended by adding further Processing and Verification Stations. Additional functionality includes language packages for recognition of Chinese/ Japanese/ Korean, Vietnamese, Hebrew and connectors for the Google Search Appliance and iFilter for Microsoft Windows Search and Microsoft Office SharePoint Server Search.*

Maintenance and Support

One year maintenance, support and updates are included in the original purchase. After this period customers have the option to purchase an extended support and maintenance agreement.

Trial Versions

ABBYY offers fully functional trial versions with time and quantity limits. More information about the above mentioned product editions, pricing, supported formats and technologies can be obtained through ABBYY or one of its partners.

* Specialised features available with Recognition Server Extended Edition only. Certain functions are available as Add-on modules to Recognition Server Extended Edition.

** Available as Add-on Modules for Recognition Server Professional Edition, included in Recognition Server Extended Edition

ABBYY®

ABBYY Europe GmbH
Eisenheimerstr. 49, 80687 Munich, Germany
Tel: +49 89 51 11 59 - 0, Fax: +49 89 51 11 59 - 59
sales_eu@abbyy.com

www.ABBYY.com