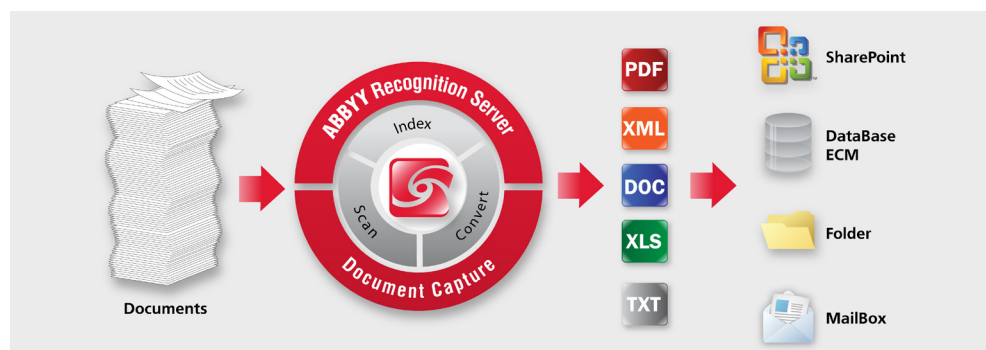


# ABBYY Recognition Server™ 3.5

## Robust OCR and Document Capture Solution

ABBYY Recognition Server 3.5 automatically converts large volumes of paper documents or document images into fully searchable electronic text for demanding e-discovery, archiving, business process automation and enterprise search systems. This scalable, robust, server-based solution can be easily installed and integrated with back-end systems and into specific business processes to improve efficiency within organizations.



### Where to use?

In order to increase productivity, organizations are leveraging cutting-edge technologies to automate processes and get the most out of the information collected every day. Recognition Server 3.5 converts information into usable data, such as converting paper-based documents into searchable formats like PDFs with indexed data attached for easier access.

### Indexing and Archiving

Convert a large number of documents into full-text searchable formats to build or optimize your existing archive. Recognition Server 3.5 offers both full-text conversion and indexing of image-based documents. Documents can be converted in PDF, XML and other formats, and critical fields can be prompted for indexing by the operator. Its easy-to-use integration tools help deliver data directly into your archive, SharePoint® or any other back-end system.

### Enterprise Search Systems

Recognition Server 3.5 with the Google Search Appliance™ (GSA) and Microsoft® iFilter modules work as a background OCR service that helps

organizations unlock and access documents saved in image-based formats. Recognition Server 3.5 gives enterprise organizations an automatic and cost-effective way to organize and access knowledge that was hidden from search results in the past.

### e-Discovery

Recognition Server 3.5 provides a flexible and accurate technology resource at the very front of the e-Discovery process. While much of e-evidence is in a searchable electronic format, email attachments of scanned PDF, JPEG or TIFF image files are not usually searchable. You can quickly and easily convert “locked” or non-searchable documents into searchable files with Bates stamping while keeping the native format retained and then pass them to searchable archives.

### Every-day Conversion

As a 24/7 OCR service, every employee in your organization can have access to Recognition Server 3.5's OCR capabilities without installation on each workstation in the local network. Simply drag and drop the desired document into the specified folder to convert and deliver accurate searchable results.

## Key Features and Benefits

### Accuracy

Over 20 years of recognition technology experience powers Recognition Server 3.5 to ensure the accurate conversion of text, even from low-quality images.

### Scalability

Parallel processing allows Recognition Server 3.5 to maximize its capacity by automatically distributing the workload. Processing performance can easily be increased by using additional workstations and CPU cores.

### Fault Tolerance

Stability and fault tolerance is ensured by integrated functions, such as job and server logging, OCR auto-start, management processes after restart and others.

### Integration

Recognition Server 3.5 can be used either as a stand-alone solution or interact with external applications. Integration scripts, XML-Ticket Support, COM-based API and Web Service API are available.

### Multilingual

Recognition Server 3.5 recognizes documents in more than 190 languages, including Chinese, Japanese, Korean, Hebrew and Thai.

# Functionality Overview

## Setup and Management

### Administration

Recognition Server 3.5 is administered remotely via the Microsoft Management Console (MMC). All system settings, including workflows, job lists, stations properties, licenses and server log files, can be edited at a central location.

### Workflow and Job Distribution

The Recognition Server workflow is the smallest administrative unit, and it includes all the processing parameters, including input sources from which the documents are taken for processing. A job can include a

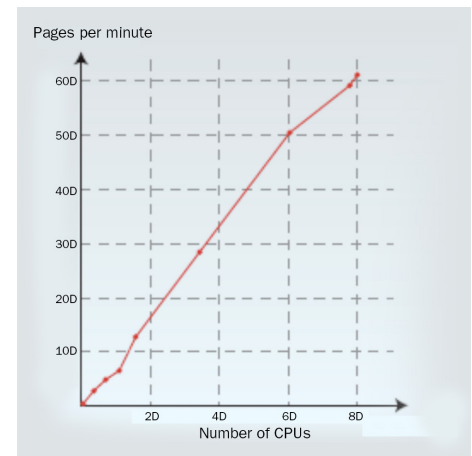
single page image, a multiple page image or a PDF file. The administrator sets priorities for each workflow in advance so that all jobs are automatically loaded throughout the entire system.

### Scheduled and Efficient Processing

Specific workflows can be established to take place at specific times or on a regular basis (i.e., daily, weekly or monthly). This kind of flexible scheduled routing increases efficiency, especially when processing is performed during evening or off-peak hours. Administrators can also configure the system so that different Processing Stations are used at different times.

## Scalability

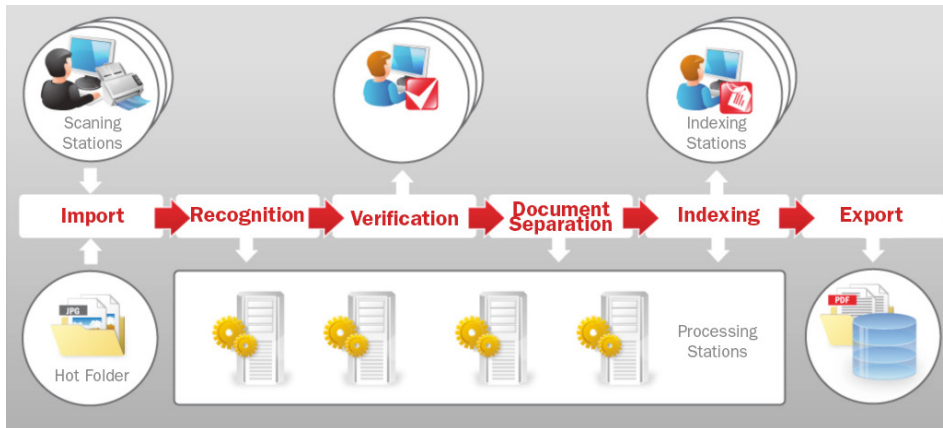
Recognition Server 3.5 can satisfy the needs of any organization's processing needs. The Server Manager and Processing Stations can be installed on a multi-core system or distributed in the network – either way, the flexible architecture will utilize all available resources for better productivity. Systems that process hundreds of pages per minute can easily be established.



### High Scalability for High-volume Demand

One Server Manager can manage dozens of connected Processing Stations and effectively distribute the workload among them – up to several hundred pages per minute.\*

\* This graph shows the speed increase according to ABBYY's internal testing results; the system performance may vary depending on the complexity and quality of the images, on the server and station hardware and the network configuration.



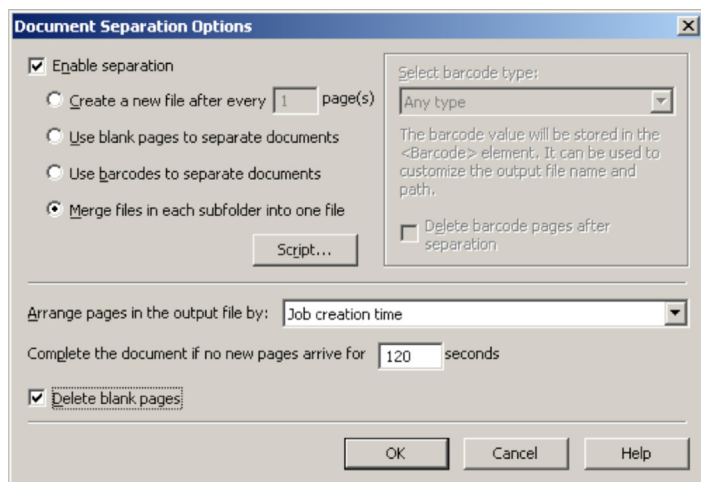
## Processing Steps:

- 1. Scanning/Import of documents.** Images can be either scanned by an operator on the Scanning Station and then sent to Recognition Server, or automatically imported by Recognition Server from an Input folder (network folder, FTP folder, or mailbox).
- 2. Recognition.** The first image file from the queue is routed to a Processing Station for recognition. If there are several Processing Stations in the system, the Server Manager will evenly distribute the files from the queue among these Processing Stations. When the Processing Station has finished processing a file, it returns the recognized file to the Server Manager and gets the next file from the queue.
- 3. Verification (optional).** If verification is turned on in the workflow settings, appropriate pages will be queued for verification after they are recognized. The Server Manager will route the queued pages to available Verification Stations, and when the verification is complete, the pages are returned to the Server Manager.
- 4. Document separation (optional).** If the document separation is enabled in Recognition Server 3.5, it is performed after recognition and verification. Documents can be separated using one of the built-in methods (by blank pages, barcodes or fixed number of pages), or by a custom rule defined in a script.
- 5. Classification and indexing (optional).** If the indexing option is on for the current workflow, the documents are then queued for indexing. Indexing can be done either automatically, by a script or by an operator by simply pointing and clicking on the Indexing Station.
- 6. Export.** In the final stage, the Server Manager delivers the output documents to their destination (which can be a network folder, a SharePoint document library, or an e-mail address). Additionally, scripts can be applied for intelligent routing and delivery of documents to ECM systems based on document properties and metadata.

# Processing Features

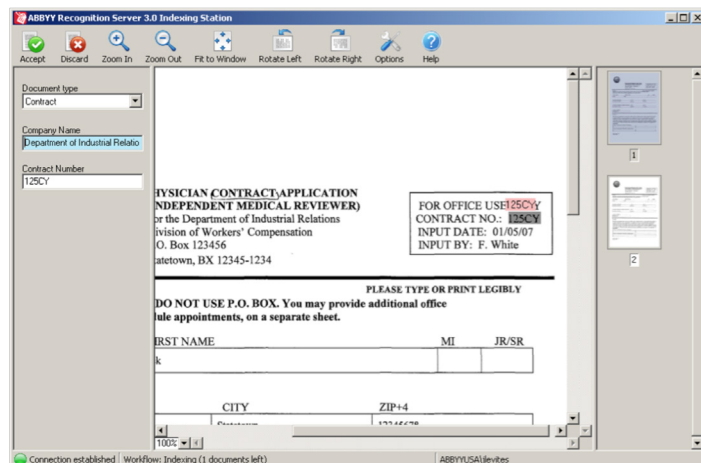
## Document Separation and Naming

Recognition Server 3.5 can separate scanned batches into different documents. They can be divided by a fixed number of pages, or by using blank pages or barcodes. Documents in a sub-folder can also be merged into a single document. Date, time, specific text and barcode values can be used for naming documents.



## Classification and Point-and-Click Indexing

Indexing can be done either automatically, by a script, or by an operator on the Indexing Station. The operator can easily select the document type and fill attributes by populating the recognized data from the image by simply clicking on the field on the image, or by typing in the data if necessary. The script can use the content of the document to set the document type, fill in attributes and decide whether the document needs to be routed to an Indexing Station for further point-and-click indexing and verification.

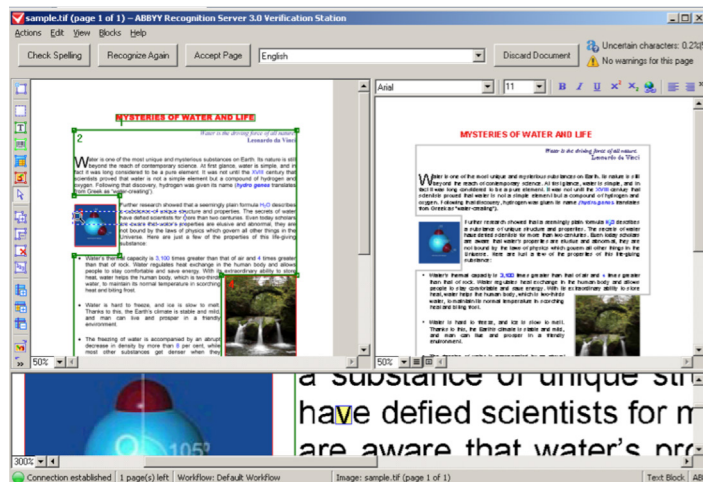


Based on ABBYY's award-winning recognition technology, Recognition Server 3.5 accurately converts documents into searchable formats in over 190 languages, including Chinese, Japanese, Korean\*, Hebrew\*, Vietnamese\* and Thai\*. Recognition Server also supports popular 1D and 2D barcodes, and provides a set of image-enhancement functions, such as splitting facing pages (for book scans), converting color images to black-and-white, correcting skew and detecting image orientation.

\* This functionality is provided as an additional module.

## Quality Control and Verification

Verification Stations enable the employees or operators to verify processing results. For example, verifiers can resize and check block types (image, text or tables) or change the OCR results. Verification Stations can be installed on multiple workstations. The Server Manager controls the number of simultaneously used stations through concurrent licensing.



## Output Formats

Recognition Server 3.5 offers a variety of different output 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, and the results can be sent by e-mail or placed in different locations, i.e., a file or SharePoint Server. Recognition Server 3.5 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 color documents.

# Integration Features

## Integration Features

Recognition Server 3.5 can easily be integrated with external applications like enterprise content management and workflow systems, expanding existing applications and automated workflows. SharePoint connector is available out-of-the-box. For integrating with other systems or implementing web services or other service-oriented applications. Recognition Server 3.5 offers several integration options, including: Scripting, XML-Ticket Support\*, COM-based API\*, Web Service API\*.

*\* This functionality is provided as an additional module.*

## Enterprise Search Systems

Recognition Server 3.5 enables complete text search capabilities for SharePoint, Microsoft Search and Google Search Appliance (GSA) systems. With Recognition Server 3.5 providing full-text recognition technologies, image-based documents like TIFFs, JPEGs, PDFs and others can be easily converted into fully searchable data for your SharePoint, MS Search and GSA. Recognition Server 3.5 gives enterprise organizations an automatic and cost-effective way to organize and access knowledge that was hidden from search results in the past.

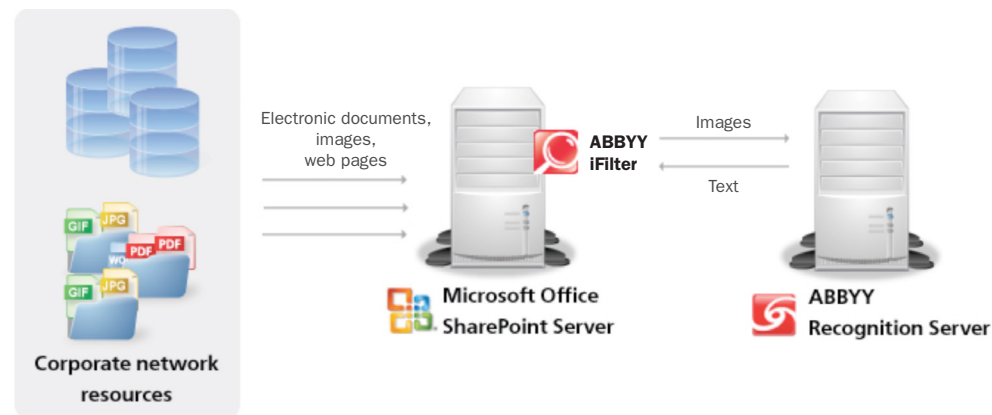
## Google Search Appliance (GSA)

Recognition Server 3.5 for the GSA runs automated and unattended, extracting text results from image-based documents and delivering results to the GSA. The system searches specified network file shares and nested subfolders, picks any image files and sends them to Recognition Server for OCR. Recognition Server 3.5 picks up the OCR results and submits the recognized document text to GSA in an XML feed with a link to the original image.



## Microsoft Office SharePoint Server and Windows Search

Graphical documents with content that cannot be indexed are sent to Recognition Server 3.5, which recognizes the documents and returns the results to Microsoft Office SharePoint Server or Windows Search for indexing, after which the documents become available for full-text search.



## Specifications

### Benefit Specs

**Input formats:** BMP, PCX, JPEG, JPEG2000, PNG, TIFF, DjVu, DCX, PDFs (including PDF/A)

**OCR:** Over 190 languages, including Chinese, Japanese, Korean, Hebrew and Thai

**Barcodes:** 1D barcodes, including EAN, Check, IATA, UCC and UPC; 2D barcodes, including PDF417, AZTEC, Data Matrix and QR Code

**Output Formats:** PDF, PDF/A, RTF, DOC, DOCX, XLS, XLSX, TXT, CSV, HTML, TIFF, JPEG, JPEG 2000, XML

### Export to Microsoft SharePoint

**Export to back-end applications via API and scripts**

**Integration:** Scripting, Web Service API, COM API, XML Tickets

**Enterprise Search modules for MS iFilter and GSA**

### System Requirements

- PC with a 500MHz or higher processor
- Operating System: Microsoft® Windows® 7, Microsoft Windows Server® 2008, Windows Vista®, Windows Server 2003, Windows XP, Windows 2000, and 64-bit versions of Windows Server 2008, Windows Vista, Windows Server 2003, Windows XP
- 128MB of RAM and an additional 100MB for each recognition process
- 700MB of hard-disk space
- Microsoft .NET Framework 2.0 or later
- Microsoft Outlook 2000 or later (for processing and sending e-mails)

**ABBYY USA**  
880 N. McCarthy Blvd.,  
Suite 220  
Milpitas, CA 95035, USA  
Tel 408.457.9777  
Fax 408.457.9778  
sales@abbyyusa.com

**www.ABBYY.com**

**Learn More at [www.ABBYY.com/recognition\\_server](http://www.ABBYY.com/recognition_server)**