

## ABBYY® Recognition Server 3.0

Professional Edition & Extended Edition



### Potente soluzione OCR e di acquisizione dei dati

ABBYY Recognition Server, una potente soluzione basata su server, consente alle aziende di convertire in modo efficiente i documenti cartacei e le immagini di documenti in file elettronici ricercabili e riutilizzabili. L'avanzata tecnologia supporta la scannerizzazione dei documenti, un OCR full text estremamente preciso e l'estrazione di metadati per ottenere contenuti adatti agli archivi elettronici, all'Enterprise Content Management (ECM), all'Enterprise Search e ad altri processi aziendali basati su documenti. Un'architettura scalabile permette l'elaborazione di grossi volumi di documenti, eseguita efficientemente e gestita in modo centralizzato. Le funzioni di acquisizione dei documenti e l'OCR automatizzato di Recognition Server possono essere integrati facilmente in sistemi back end e in altre applicazioni.

### Dove è impiegato ABBYY Recognition Server?

Oggigiorno il successo aziendale dipende molto dall'incremento continuo della produttività e dall'utilizzo ottimale delle informazioni con cui le imprese vengono sommerse quotidianamente. Utilizzando tecnologie innovative, Recognition Server automatizza i processi di elaborazione dei documenti e converte le informazioni sia cartacee, sia digitali in formati riutilizzabili, come ad esempio i PDF indicizzati. I file ottenuti sono così ricercabili e possono essere utilizzati in altri processi aziendali.

- **Archiviazione dei documenti e conversione di documenti e dati già archiviati in precedenza**

Recognition Server 3.0 permette la conversione di un grosso numero di documenti in formati full text ricercabili per la gestione e l'archiviazione dei documenti. Offre sia la conversione full text, sia la semplice indicizzazione dei documenti basati su immagini. Dopo aver convertito i documenti in formato PDF, XML\* e altri, gli operatori hanno la possibilità di controllare manualmente la qualità del riconoscimento o i campi dell'indice. I tool per l'integrazione di questa soluzione, facili da utilizzare, sono di grande aiuto nell'inserimento dei dati direttamente in archivi, SharePoint server o in altri sistemi back end.

- **Conversione ventiquattrore su ventiquattro, sette giorni su sette**

In quanto sistema basato su server, Recognition Server offre l'accesso costante alla conversione OCR, senza il bisogno di installare software sulle singole stazioni di lavoro. Gli utenti possono semplicemente inviare una e-mail, oppure trascinare i documenti in una cartella predefinita per ottenere in modo rapido informazioni e documenti esatti e ricercabili. Inoltre, Recognition Server può essere collegato a dispositivi MFP e scanner di rete. In questo modo, i documenti scannerizzati o copiati possono essere elaborati automaticamente tutti i giorni.

- **Incremento della produttività nel settore giuridico**

Recognition Server 3.0 offre una tecnologia flessibile e precisa per la gestione e la catalogazione delle informazioni. Un'accurata gestione dei contenuti è particolarmente importante per gli studi legali, dove quotidianamente affluisce e viene elaborata una grossa quantità di informazioni. Anche se la maggior parte dei documenti elettronici sono inviati in un formato ricercabile, gli allegati e-mail sono spesso in formati non editabili, come ad esempio PDF scannerizzati, JPEG o TIFF, che rischiano di rallentare il processo di rilevamento dei dati importanti. Recognition Server converte in modo rapido e semplice queste informazioni "non accessibili" in file ricercabili (completi di bates stamp e nel formato originale).

- **Sistemi di Enterprise Search**

Recognition Server 3.0 offre dei moduli per i sistemi di Enterprise Search. I moduli Google Search Appliance™ (GSA) e Microsoft® IFilter\*\* funzionano come servizio OCR in background per supportare le aziende nell'acquisire le informazioni da formati basati su immagini. In questo modo, le imprese hanno la possibilità di accedere automaticamente e in maniera economica alle informazioni che fino ad ora non potevano essere ritrovate con la funzione di ricerca.

### OCR e conversione di PDF tramite un'architettura server-based

- **Automatizzato e indipendente.** OCR programmato per iniziare l'elaborazione dopo l'arrivo dei documenti nelle "cartelle sorvegliate", oppure per processare i documenti a un orario prestabilito. Funziona come processo in background e indipendentemente dalle stazioni di lavoro client.
- **Punti di accesso multipli.** Accesso ai documenti da dispositivi MFP, da scanner di gruppi di lavoro, dalla rete, da cartelle FTP e dalla posta in entrata.
- **Elaborazione di grossi volumi di documenti.** Massimizzazione delle capacità tramite l'elaborazione parallela e la distribuzione automatica del carico di lavoro.
- **Scalabilità.** Innalzamento del rendimento dell'elaborazione grazie all'aggiunta di ulteriori stazioni di lavoro e core della CPU.
- **Gestione centralizzata.** Gestione e amministrazione centralizzate.
- **Precisione.** Oltre 20 anni di esperienza nel campo della tecnologia di riconoscimento offrono come risultato una conversione del testo precisa e accurata.
- **Internazionale.** Riconoscimento di più di 190 lingue.
- **Tolleranza agli errori.** Protocollo eventi, avvio automatico e riavvio della conversione OCR e inoltro dei documenti quando necessario.
- **Integrazione flessibile.** Il supporto ticket XML, l'API basata su COM e l'API dei servizi Web consentono l'integrazione in altre applicazioni.

\* Funzioni specializzate disponibili solo con Recognition Server Extended Edition. Alcune funzionalità sono disponibili solo come moduli add-on per l'Extended Edition.

\*\* Disponibili come moduli add-on per Recognition Server Professional Edition, inclusi in Recognition Server Extended Edition.

# Panoramica delle funzionalità



## Le fasi dell'elaborazione dei documenti

### 1. Scannerizzazione e importazione dei documenti:

- **Scannerizzazione:** le immagini sono scannerizzate da un operatore sulla Stazione di scannerizzazione e poi inviate a Recognition Server.
- **Importazione dei documenti:** il Server Manager importa i file dalla sorgente di input designata (ad es. una cartella condivisa, una cartella FTP o una cartella di posta elettronica) e li inserisce nella coda di elaborazione.

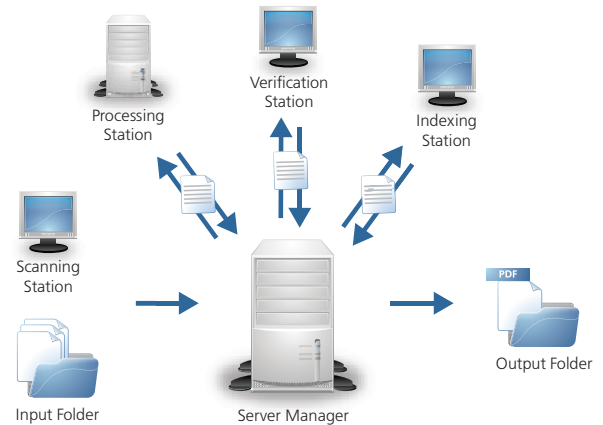
**2. Riconoscimento:** il sistema distribuisce i file uniformemente tra le Stazioni di elaborazione disponibili per il riconoscimento. Ogni workflow per l'elaborazione dei documenti sulle Stazioni di elaborazione disponibili è amministrato e monitorato dal Server Manager. Entrambe le componenti funzionano come servizi di Windows® e possono essere installate separatamente in rete, oppure su una stazione di lavoro.

**3. Verifica (facoltativa):** in caso di scarsa qualità del documento di input, oppure quando è richiesta una precisione elevata, i documenti possono essere inviati facoltativamente ad una Stazione di verifica per il controllo manuale. Il Server Manager indirizza i documenti verso le Stazioni di verifica, a seconda delle configurazioni del workflow\*\*\*.

**4. Separazione dei documenti:** dopo il riconoscimento e la verifica, i documenti possono essere suddivisi utilizzando pagine bianche, codici a barre, oppure in base a numeri fissi di pagine. Inoltre, è possibile utilizzare degli script per creare delle regole di separazione personalizzate.

**5. Classificazione e indicizzazione (facoltative):** le imprese possono scegliere se i dati debbano essere indicizzati manualmente nella Stazione di indicizzazione, oppure in modo automatico utilizzando script personalizzati che permettono l'esportazione e la consegna intelligenti, basate sulle proprietà dei documenti e sui metadati.

**6. Esportazione:** dopo il riconoscimento e la verifica, il Server Manager invia il documento di output alla sua destinazione che può essere una cartella di rete, una libreria SharePoint®\*, un indirizzo e-mail, oppure delle applicazioni che utilizzano l'API di Recognition Server.



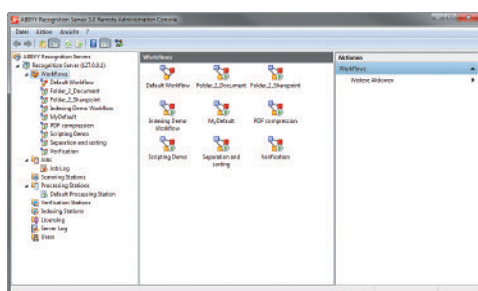
## Configurazione e gestione

### Gestione e amministrazione

ABBYY Recognition Server è amministrato in modalità remota tramite la Microsoft Management Console (MMC). Tutte le impostazioni del sistema, inclusi i workflow, le liste dei compiti, le proprietà della Stazione di elaborazione, le licenze e i file dei protocolli del server, possono essere elaborati da una postazione centrale.

### Workflow di conversione e distribuzione dei compiti

Un workflow di Recognition Server consiste in una serie di parametri di elaborazione che includono la sorgente di input del documento, le fasi dell'elaborazione e i metodi di consegna. Un compito è la più piccola unità di elaborazione e può contenere una o più immagini, oppure dei file PDF. L'amministratore stabilisce in anticipo delle priorità per ogni workflow in modo da assicurare che tutti i compiti siano processati in ordine di importanza.

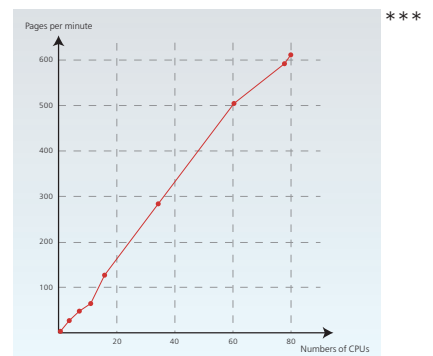


### Elaborazione programmata ed efficiente

Gli amministratori hanno la flessibilità di configurare particolari workflow perché siano eseguiti in orari prestabiliti, oppure regolarmente (ad esempio giornalmente, settimanalmente o mensilmente), in modo da incrementare l'efficienza e sfruttare i periodi di minor flusso di lavoro (ad esempio durante le ore notturne). Gli amministratori possono, inoltre, configurare il sistema per utilizzare differenti Stazioni di elaborazione in momenti diversi.

### Scalabilità

ABBYY Recognition Server è in grado di soddisfare le necessità delle imprese di ogni dimensione e con differenti esigenze riguardanti l'elaborazione dei documenti. Il Server Manager e le Stazioni di elaborazione possono essere installati su un sistema multicore, oppure distribuiti in rete. In entrambi i casi, l'architettura flessibile è scalabile in modo efficiente e permette agli amministratori di predisporre facilmente i sistemi che hanno la capacità di elaborare centinaia di pagine al minuto.





# Panoramica delle funzionalità

## Funzionalità

### Input e scannerizzazione

ABBYY Recognition Server offre un'applicazione per la scannerizzazione diretta. La Stazione di scannerizzazione supporta le interfacce di scansione TWAIN, WIA e ISIS e offre tool per il controllo della qualità visiva dei documenti prima del riconoscimento. I tool per la pre-elaborazione che sono disponibili con la Stazione di scannerizzazione includono l'anteprima del documento, diverse opzioni per il miglioramento dell'immagine, la possibilità di definire il tipo di batch, la rimozione manuale dei dati sensibili e l'ordine della successione delle pagine.

### Riconoscimento e conversione dei PDF

Recognition Server si basa sulla pluripremiata tecnologia di riconoscimento di ABBYY, celebre per la sua precisione e per la sua stabilità. Il sistema supporta una vasta gamma di funzioni di riconoscimento e di formati di input, tra cui:

- **pre-elaborazione delle immagini:** divisione delle pagine doppie (ad es. per la scansione di libri), conversione di immagini a colori o in scala di grigi in immagini in bianco e nero, rimozione del rumore di sottofondo e altro ancora;
- **tipo di stampa:** testo normale, testo battuto a macchina, stampante a matrice di punti, OCR-A, OCR-B, e MICR (E13b);
- **lingue:** più di 190 lingue, elaborazione accurata anche di documenti multilingue\*;
- **caratteri speciali:** supporto per Fraktur, Schwabacher e molti altri caratteri gotici stampati tra il 1700 e il 1937 in inglese, tedesco, francese, italiano e spagnolo\*;
- **codici a barre:** riconoscimento dei codici a barre 1D e 2D più comuni, inclusi i codici 2D Aztec, Data Matrix e QR Code, disposti in qualsiasi angolo del documento e in qualsiasi posizione;
- **3 modalità differenti di elaborazione:** precisa, veloce e bilanciata.

### Personalizzazione tramite scripting

Grazie all'utilizzo di script, gli utenti hanno la possibilità di personalizzare i processi interni, la gestione delle eccezioni e l'esportazione dei documenti. Possono essere usati diversi tipi di script per creare regole di separazione più flessibili, automatizzare l'indicizzazione semplice, aggiungere funzioni di export dinamiche e personalizzate e semplificare la pubblicazione dei documenti di output su sistemi esterni. È possibile utilizzare gli script per gestire le eccezioni, al fine di elaborare separatamente i documenti che richiedono speciale attenzione.

### Controllo della qualità e verifica\*\*\*\*

Le Stazioni di verifica, che possono essere installate su stazioni di lavoro multiple e controllate utilizzando una licenza concomitante, permettono agli operatori di controllare o verificare manualmente i risultati dell'elaborazione. I verificatori hanno la possibilità di controllare la ricostruzione del layout (immagine, testo, tabelle), editare il testo riconosciuto e controllare l'ortografia.

### Separazione e denominazione dei documenti

Recognition Server può utilizzare pagine vuote o codici a barre come separatori di documenti. Inoltre, è in grado di creare documenti con un

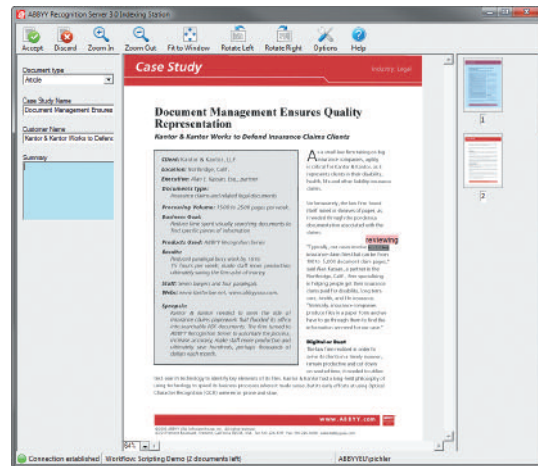
numero fisso di pagine e permette di unire i file contenuti in una sottocartella in un unico documento. Gli utenti hanno anche la possibilità di scegliere di creare convenzioni per la denominazione automatica dei documenti utilizzando caratteristiche come data, ora, testo specifico e valori dei codici a barre. Inoltre, gli amministratori possono creare regole personalizzate per la separazione dei documenti tramite script.

### Classificazione e indicizzazione "point-and-click"

La Stazione di indicizzazione offre un'interfaccia che permette agli operatori di specificare il tipo di documento e i suoi attributi. Sono disponibili tre modi differenti per definire gli indici:

- **indicizzazione "type-in":** gli operatori scrivono negli indici manualmente;
- **indicizzazione "point-and-click":** gli operatori estraggono le keyword per gli indici cliccando con il mouse sull'immagine di un documento
- **indicizzazione per frasi chiave:** gli operatori utilizzano una funzione "elastico" per tracciare il contorno di frasi, linee multiple o paragrafi.

Inoltre, gli amministratori possono creare e definire funzioni automatiche di indicizzazione basate sulla struttura del documento e sul suo contenuto, utilizzando degli script.



### Formati di output

Recognition Server 3.0 offre una varietà di formati di output, tra cui DOC, DOCX, RTF, XML\*, XLS, XLSX, HTML, TIFF, JPEG, JPEG2000 e altri ancora. È possibile generare formati multipli di output per un singolo documento di input. Recognition Server supporta l'export in "semplici" PDF ricercabili, PDF lineari e tagged, file PDF con opzioni di sicurezza e codifica e PDF/A per l'archiviazione a lungo termine. Inoltre, grazie alla tecnologia MRC, il sistema permette di generare PDF altamente compressi, adatti all'archiviazione dei documenti a colori.

### Esportazione

I risultati del riconoscimento e della conversione possono essere inviati via e-mail, oppure depositati in postazioni differenti come sistemi ECM, archivi o Sharepoint Server\*.

\* Funzioni specializzate disponibili solo con Recognition Server Extended Edition. Alcune funzionalità sono disponibili solo come moduli add-on per l'Extended Edition.  
 \*\* Disponibili come moduli add-on per ciascuna Edition.  
 \*\*\* Il grafico e le informazioni riguardanti la capacità di elaborazione si basano su test interni. Il rendimento del sistema può variare a seconda della qualità delle immagini, del rendimento dell'hardware, dalla configurazione di rete e da altri fattori.  
 \*\*\*\* La disponibilità della Stazione di verifica è prevista per il 4° trimestre del 2010.

# Integrazione e disponibilità

## Requisiti generali di sistema:

- **PC** con processore Intel® Core™2/2 Quad/Pentium® / Celeron® / Xeon™, AMD K6/ Turion™/ Athlon™/ Duron™/ Sempron™ o compatibile, con una velocità di clock di almeno 2 GHz.
- **Sistema operativo:** Microsoft® Windows® 7, Windows Server® 2008, Windows Vista®, Windows Server 2003, Windows XP
- **Memoria:** Server Manager: 1 GB di RAM; Stazione di scannerizzazione: 1 GB di RAM; Stazione di elaborazione: 512 MB di RAM più 300 MB addizionali per ogni processo di riconoscimento; Stazione di indicizzazione: 768 MB di RAM.
- **Spazio libero sull'hard disk:** Server Manager: 20 MB per l'installazione, 1 GB addizionale per le operazioni di programma; Stazione di scannerizzazione: 1 GB; Stazione di elaborazione: 600 MB per l'installazione e 1 GB addizionale per le operazioni di programma; Stazione di indicizzazione: 500 MB per l'installazione e 1 GB per le operazioni di programma.
- **Microsoft .NET Framework 2.0** o superiore per il salvataggio dei file in Microsoft SharePoint Server.
- **Microsoft Outlook 2000** o più recente per l'elaborazione dei messaggi e-mail tramite Microsoft Exchange Server.
- **Microsoft IIS 5.1** o versione successiva (per la Web API).
- **Scanner** con supporto TWAIN, WIA o ISIS.
- **Scheda video e monitor** con risoluzione 1024 x 768.

Lo spazio richiesto sull'hard disk per le operazioni di programma può essere maggiore a seconda della complessità, della qualità e del numero delle immagini. I requisiti di sistema possono variare a seconda del componente server o dei moduli aggiuntivi utilizzati. Per informazioni più dettagliate contattare il rappresentante delle vendite ABBYY.

## Formati di input:

- BMP
- PCX
- JPEG
- JPEG 2000
- PNG
- TIFF
- PDF (fino a PDF 1.6)
- DjVu

## Lingue OCR:

più di 190 lingue, tra cui cinese, giapponese, coreano ed ebraico.

## Formati di output:

- DOC (Microsoft Word)
- DOCX (Microsoft Word 2007)
- RTF
- PDF, PDF/A, PDF (Versione 1.6)
- HTML
- CSV (comma separated values)
- TXT
- XLS (Microsoft Excel)
- XLSX (Microsoft Excel 2007)
- file immagine originali
- TIFF
- JPEG
- JPEG2000

## Integrazione

ABBYY Recognition Server può essere facilmente connesso ad applicazioni esterne, inclusi i sistemi ECM (Enterprise Content Management) e i sistemi per la gestione del workflow, al fine di espandere le capacità delle applicazioni esistenti e di automatizzare la conversione e la consegna dei documenti. Le aziende che sono interessate a implementare servizi o altre applicazioni service-oriented possono facilmente integrare Recognition Server.

ABBYY Recognition Server offre una serie di opzioni per l'integrazione avanzata. Le funzionalità includono:

- Supporto ticket XML\*
- API basata su COM\*
- Web Service API\*
- Microsoft SharePoint Export\*

## Connettori per Enterprise Search

ABBYY Recognition Server 3.0 espande le capacità dei sistemi di ricerca di Google e Microsoft grazie ad una connessione a Google Search Appliance, Microsoft Office SharePoint Server e Windows Desktop Search. Questi moduli add-on permettono l'indicizzazione del contenuto full text di fax, accordi scannerizzati, corrispondenza e qualsiasi altro documento in formato TIFF, JPEG o PDF\*\*.

## Licenze

Recognition Server è disponibile in due diverse versioni:

### Recognition Server 3.0 Professional Edition

Recognition Server 3.0 Professional Edition offre funzionalità standard per le imprese che desiderano un servizio per la di conversione OCR automatizzata da eseguirsi in background e che necessitano di un adattamento parziale e quasi nessuna integrazione con altre applicazioni.

### Recognition Server 3.0 Extended Edition

Recognition Server 3.0 Extended Edition offre un'integrazione completa con applicazioni esterne e permette l'implementazione come parte di un'architettura Web Service. Le funzionalità aggiuntive offerte nell'Extended Edition includono: lingue OCR addizionali (tailandese ed ebraico), esportazione XML, Web Service API, API basata su COM, supporto ticket XML e la possibilità di esportare i file in Microsoft SharePoint. FineReader XIX (riconoscimento di caratteri gotici) e l'OCR per cinese, giapponese e coreano sono disponibili solo con Recognition Server 3.0 Extended Edition.

\* Funzioni specializzate disponibili solo con Recognition Server Extended Edition. Alcune funzionalità sono disponibili solo come moduli add-on per l'Extended Edition.

\*\* Disponibili come moduli add-on per ciascuna Edition.

### Moduli add-on

Entrambe le versioni possono essere estese aggiungendo ulteriori Stazioni di elaborazione e di verifica. Le funzionalità aggiuntive includono i pacchetti linguistici per il riconoscimento di cinese/giapponese/ coreano, vietnamita ed ebraico, così come connettori per Google Search Appliance, iFilter per Microsoft Windows Search e Microsoft Office SharePoint Server Search.\*

### Manutenzione e assistenza

Con il primo acquisto di una licenza sono inclusi un anno di manutenzione, così come assistenza ed aggiornamento. Trascorso questo periodo, i clienti hanno la possibilità di estendere il supporto attraverso contratti di servizi e assistenza aggiuntivi.

### Versioni di prova

ABBYY mette a disposizione versioni di prova aventi tutte le funzionalità, che permettono la valutazione del prodotto per un periodo ed una quantità limitati. Maggiori informazioni riguardanti le versioni del prodotto, i prezzi, i formati supportati e le tecnologie possono essere ottenute tramite ABBYY o uno dei suoi partner.

# ABBYY®

ABBYY Europe GmbH  
Eisenheimerstr. 49, Monaco di Baviera, Germania  
Tel: +49 89 5111159 - 0, Fax: +49 89 5111159 - 59  
sales\_eu@abbyy.com

[www.ABBYY.com](http://www.ABBYY.com)