

## ABBYY® Recognition Server 3.5

*Professional Edition & Extended Edition*



### Potente soluzione OCR e di acquisizione documenti

ABBYY Recognition Server, una potente soluzione basata su server, permette 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 scansione dei documenti, l'OCR full text estremamente preciso e l'estrazione dei metadati per creare contenuti adatti ad archivi elettronici, ai sistemi di gestione dei contenuti (ECM), all'Enterprise Search e ad altri processi aziendali che si basano su documenti. Un'architettura scalabile permette l'elaborazione di volumi elevati, eseguita efficientemente e gestita in modo centralizzato. Le funzioni di acquisizione dei documenti e di OCR automatizzato offerte da Recognition Server possono essere integrate facilmente in sistemi di back end e in altre applicazioni.

### Dove si usa ABBYY Recognition Server?

Oggi, il successo aziendale dipende molto dall'incremento continuo della produttività e dall'utilizzo ottimale delle informazioni con cui le imprese sono 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 PDF indicizzati. I file ottenuti sono così ricercabili e possono essere usati in altri processi aziendali.

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

Recognition Server permette la conversione di volumi elevati di documenti in formati full text ricercabili per la gestione e l'archiviazione dei documenti stessi. Offre sia la conversione full text, sia la semplice indicizzazione dei documenti basati su immagini. Dopo aver convertito i documenti in PDF, XML\* o altri formati, gli operatori hanno la possibilità di controllare la qualità del riconoscimento e di inserire manualmente i campi dell'indice. Grazie ai tool di facile utilizzo per l'integrazione della soluzione, è possibile inserire in modo automatico i dati direttamente in archivi, in SharePoint Server o in altri sistemi di back end.

- **Conversione dei documenti sempre disponibile**

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

- **Incremento della produttività nel settore giuridico**

Recognition Server offre una tecnologia flessibile e precisa per gestire e catalogare il flusso di informazioni in entrata ogni giorno in uno studio legale. Anche se la maggior parte dei documenti elettronici è disponibile in formato ricercabile, spesso gli allegati e-mail quali PDF scannerizzati, JPEG o TIFF non lo sono e rischiano quindi di rallentare l'acquisizione dei dati rilevanti. Recognition Server converte in modo rapido e semplice queste informazioni "non accessibili" in file ricercabili (completi di bates stamp e mantenendo la formattazione originale).

- **Sistemi di Enterprise Search**

Recognition Server offre dei moduli per i sistemi di Enterprise Search. I moduli per Google Search Appliance™ (GSA) e Microsoft® IFilter\*\* supportano le aziende nell'acquisizione delle informazioni da documenti in formato immagine. In questo modo, le imprese hanno la possibilità di accedere automaticamente e in maniera economica alle informazioni disponibili nella rete aziendale 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 avviare l'elaborazione appena i documenti arrivano nelle "cartelle sorvegliate", oppure per processare i documenti ad 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 casella di posta in arrivo.

- **Elaborazione di volumi elevati di documenti.**

Massimizzazione delle capacità tramite l'elaborazione parallela e la distribuzione automatica del carico di lavoro.

- **Scalabilità.** Incremento 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 estremamente accurata.

- **Internazionale.** Riconoscimento in più di 190 lingue.

- **Tolleranza agli errori.**

Protocollo eventi, avvio automatico e riavvio del processo OCR e routing 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.

\* L'esportazione in XML è disponibile solo in Recognition Server Extended Edition.

\*\* Il connettore a Google Search Appliance (GSA) e l'IFilter per Microsoft Windows® Search (desktop e server) sono disponibili in Recognition Server Extended Edition. È possibile acquistarli come moduli add-on per Recognition Server Professional Edition.

# Panoramica delle funzionalità



## Fasi di elaborazione dei documenti

### 1. Scansione e importazione dei documenti:

- **Scansione:** le immagini sono scannerizzate da un operatore nella Stazione di digitalizzazione 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 di elaborazione dei documenti nelle 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 sulla medesima 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 essere controllati in modalità manuale da un operatore. 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 divisi utilizzando pagine bianche, codici a barre, oppure in base ad un numero fisso 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). È possibile accedere a tutte le impostazioni del sistema, inclusi i flussi di lavoro, le liste dei lavori, le proprietà della Stazione di elaborazione, le licenze e i file dei protocolli del server da una postazione centrale.

### Workflow di conversione e distribuzione dei lavori

Un flusso di lavoro (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 "lavoro" 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 eseguiti in ordine d'importanza.

### Elaborazione programmata ed efficiente

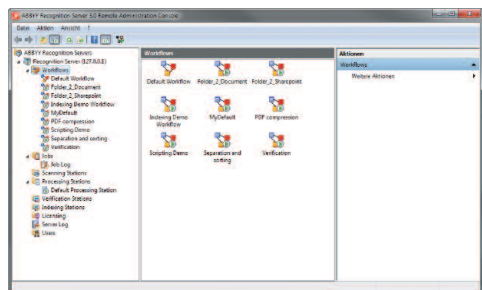
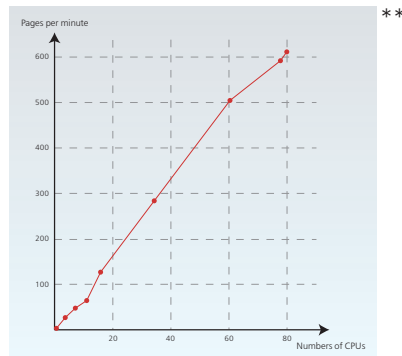
Gli amministratori hanno la flessibilità di configurare workflow particolari perché siano eseguiti in orari prestabiliti, oppure regolarmente (ad esempio giornalmente, settimanalmente o mensilmente), in modo da incrementare l'efficienza e sfruttare i periodi di minor flusso di lavoro (ad esempio durante 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 nell'ambito dell'elaborazione dei documenti. Il Server Manager e le Stazioni di elaborazione possono essere installati su un sistema multi-core, oppure distribuiti in rete. In entrambi i casi, l'architettura flessibile è scalabile in modo efficiente e permette agli amministratori di predisporre facilmente sistemi che hanno la capacità di elaborare centinaia di pagine al minuto.

\* L'esportazione verso Microsoft SharePoint è disponibile solo in Recognition Server Extended Edition.

\*\* Le informazioni e il grafico riguardanti la capacità di elaborazione si basano su test interni. Il rendimento del sistema può variare a seconda della qualità delle immagini, della prestazione dell'hardware, della configurazione di rete e altri fattori.





## Panoramica delle funzionalità

### Funzionalità

#### Importazione e scansione

Recognition Server è fornito di una Stazione di digitalizzazione per la scansione diretta. Questa stazione supporta le interfacce di scansione TWAIN, WIA e ISIS e offre altresì tool per la pre-elaborazione delle immagini dei documenti prima del riconoscimento, quali anteprima, diverse opzioni per il miglioramento dell'immagine, definizione del tipo di batch, rimozione manuale dei dati riservati e selezione dell'ordine delle pagine. I comandi di scripting permettono di regolare i processi di input, ad esempio eseguendo la separazione automatica delle pagine o il riordino delle pagine di grosse dimensioni dopo la scansione duplex.

#### 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.

- **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" 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:** riconoscimento di 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 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 amministratori hanno la possibilità di personalizzare l'elaborazione interna, la gestione delle eccezioni e l'esportazione dei documenti. È quindi possibile l'utilizzo di diversi tipi di script per creare regole più flessibili per l'import o la separazione dei documenti, così come per automatizzare l'indicizzazione o per aggiungere funzioni di export personalizzate. In questo modo, si semplifica la pubblicazione dei documenti di output in sistemi esterni. Gli script consentono altresì la definizione delle 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 diverse stazioni di lavoro e condivise da più utenti utilizzando una licenza concomitante, permettono agli operatori di verificare in modalità manuale i risultati dell'elaborazione. Gli addetti alla verifica controllano la ricostruzione del layout (immagini, testi, tabelle), modificano il testo e ne controllano l'ortografia.

#### Separazione e denominazione dei documenti

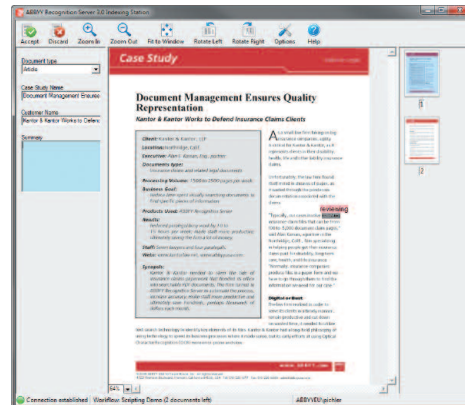
Recognition Server è in grado di utilizzare pagine vuote o codici a barre per la separazione dei documenti, oppure di separarli tramite l'impostazione di un numero fisso di pagine. I documenti contenuti in una sottocartella possono inoltre essere assemblati in un unico documento. Gli utenti hanno altresì la possibilità di creare attributi per denominare automaticamente i documenti, ad esempio tramite data, ora, testo specifico e valori dei codici a barre. Inoltre, gli amministratori possono creare tramite scripting regole personalizzate per la separazione dei documenti.

#### 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 digitano gli indici manualmente
- **Indicizzazione "point-and-click":** gli operatori estraggono le parole chiave per gli indici dall'immagine di un documento, cliccando e trascinandole con il mouse
- **Indicizzazione per frasi chiave:** gli operatori utilizzano una funzione "elastico" per marcare frasi, linee multiple o paragrafi

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



#### Formati di output

Recognition Server supporta diversi formati di output, tra cui DOC, DOCX, RTF, XML\*\*\*, XLS, XLSX, HTML, TIFF, JPEG, JPEG2000 e altri ancora. È possibile generare formati di output multipli per un singolo documento. Recognition Server supporta altresì l'esportazione verso "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-MRC altamente compressi, adatti in particolare all'archiviazione di documenti a colori.

#### Esportazione

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

\* Recognition Server Extended Edition offre più lingue di riconoscimento rispetto alla Professional Edition, includendo anche ebraico e thailandese. È inoltre possibile acquistare moduli aggiuntivi per il riconoscimento di documenti scritti in cinese, coreano, giapponese e vietnamita.

\*\* Il supporto per caratteri antichi è disponibile come modulo add-on per Recognition Server Extended Edition.

\*\*\* L'esportazione verso XML è disponibile solo in Recognition Server Extended Edition.

\*\*\*\* L'esportazione verso SharePoint Server è disponibile solo in Recognition Server Extended Edition.



# 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 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 digitalizzazione: 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 digitalizzazione: 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 versione successiva per il salvataggio dei file in Microsoft SharePoint Server
- **Microsoft Outlook 2000** o versione successiva 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 ABBYY.

## Formati di input

- BMP, PCX, DCX, GIF, TIFF, WDP
- JPEG, JPEG 2000, JBIG2, PNG
- PDF (fino alla versione 1.7), DjVu

## Lingue OCR

Supporto di più di 190 lingue\*\*\*

## Formati di output

- RTF, TXT, HTML, CSV, XML, EPUB
- DOC, DOCX (MS Word 2007)
- XLS, XLSX (MS Excel 2007)
- ALTO XML

## Formati ricercabili:

- PDF (fino alla versione 1.7); PDF/A

## Formati immagine:

- PDF sola immagine
- JPEG, JPEG 2000, JBIG2, TIFF

## 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 e di automatizzare la conversione e la consegna dei documenti. Le aziende che sono interessate a implementare servizi Web o altre applicazioni service-oriented possono utilizzare Recognition Server.

Recognition Server offre una serie di opzioni per l'integrazione in applicazioni esterne. Le funzionalità includono:

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

## Connettori per Enterprise Search\*\*

ABBYY Recognition Server è in grado di espandere le capacità dei sistemi di ricerca di Google e Microsoft, grazie al connettore per Google Search Appliance e all'iFilter per Microsoft SharePoint Server e Windows Desktop Search. Questi moduli add-on consentono l'indicizzazione del contenuto di fax, contratti scannerizzati, corrispondenza e qualsiasi altro documento in formato TIFF, JPEG o PDF.

## Licenze

Recognition Server è disponibile in due diverse versioni.

### Recognition Server 3.5 Professional Edition

La Professional Edition offre funzionalità standard per le imprese che desiderano un servizio OCR automatizzato, da eseguirsi in background e senza integrazione in altre applicazioni.

### Recognition Server 3.5 Extended Edition

L'Extended Edition offre un'integrazione completa in applicazioni esterne, come parte di un'architettura Web service. Le funzionalità aggiuntive offerte da questa versione includono: lingue OCR addizionali (tailandese ed ebraico), esportazione XML, Web Service API, API basata su COM, supporto ticket XML ed esportazione in Microsoft SharePoint. FineReader XIX (per il riconoscimento di caratteri gotici) e l'OCR per cinese, giapponese e coreano sono disponibili come moduli aggiuntivi.

### Moduli add-on

Entrambe le Edition possono essere estese aggiungendo altre stazioni, oppure ulteriori core della CPU. Le funzionalità possono inoltre essere ampliate grazie a pacchetti linguistici aggiuntivi (per il riconoscimento di arabo, cinese, giapponese, coreano, vietnamita ed ebraico), come anche a connettori per Google Search Appliance o iFilter per Microsoft Windows Search e Microsoft Office SharePoint Server Search\*\*.

### Manutenzione e assistenza

Con la manutenzione e l'assistenza, il cliente riceve regolarmente aggiornamenti del software, l'upgrade alla versione successiva e l'assistenza tecnica. L'acquisto del pacchetto di manutenzione e assistenza è obbligatorio e viene fatturato a parte a scadenza annuale.

### Versioni di prova

ABBYY mette a disposizione versioni di prova aventi tutte le funzionalità. Maggiori informazioni possono essere ottenute tramite ABBYY o uno dei suoi partner.

**Nota:** per dettagli riguardanti prezzi, funzionalità disponibili e possibilità di espansione della Professional ed Extended Edition, consultare il listino prezzi di ABBYY Recognition Server.

\* Queste opzioni di integrazione sono disponibili in Recognition Server Extended Edition.

\*\* I connettori a Google Search Appliance (GSA) e l'iFilter per Microsoft Windows Search (desktop e server) sono disponibili in Recognition Server Extended Edition. È possibile acquistarli come moduli add-on per Recognition Server Professional Edition.

\*\*\* Recognition Server Extended Edition offre un numero maggiore di lingue di riconoscimento rispetto alla Professional Edition, includendo l'ebraico e il thailandese. È inoltre possibile aggiungere ulteriori lingue come arabo, cinese, coreano, giapponese e vietnamita, acquistando i moduli add-on corrispondenti.

**ABBYY®**

**ABBYY Europe GmbH**  
Eisenheimerstr. 49 80687  
Monaco di Baviera, Germania  
Tel: +49 89 511 159 0  
sales\_eu@abby.com  
[www.ABBYY.com](http://www.ABBYY.com)

**ABBYY UK**  
sales\_UK@abby.com

**ABBYY France**  
sales\_france@abby.com

**ABBYY Spain**  
sales\_es@abby.com

**ABBYY Italy**  
sales\_it@abby.com

**ABBYY Benelux**  
sales\_benelux@abby.com

**ABBYY Scandinavia**  
sales\_nordic@abby.com