

# ABBYY Recognition Server

## Informations produit

### Solution robuste de capture de documents et de conversion en PDF

ABBYY Recognition Server est une solution serveur puissante de reconnaissance optique des caractères (OCR) pour la capture automatisée de documents et leur conversion en PDF. Elle permet aux organisations et aux prestataires de services de numérisation de mettre en place des processus rentables de conversion des documents papier et d'images en formats numériques, adaptés à l'archivage numérique à long terme et aux recherches plein texte.

ABBYY Recognition Server acquiert automatiquement les images de documents à partir de scanners, fichiers, serveurs de fax et de messagerie, mais aussi à partir des bibliothèques Microsoft® SharePoint®, et effectue une reconnaissance optique de caractères pour extraire des informations en texte intégral tout en offrant la possibilité d'ajouter des métadonnées. Les résultats sont transférés directement vers des dossiers réseau, des bibliothèques SharePoint ou d'autres systèmes de stockage et de gestion de fichiers, comme les fichiers PDF ou PDF/A interrogeables et compressés en MRC, données XML, Microsoft Word, les fichiers Excel® ou le texte brut.

Cette solution hautement évolutive vous permet de convertir de grandes quantités de documents rapidement en un laps de temps très réduit. Son déploiement rapide, son administration facile et ses routines opérationnelles automatisées confèrent à ABBYY Recognition Server un retour sur investissement rapide.

### Comment tirer les bénéfices d'ABBYY Recognition Server ?

La possibilité de convertir des documents commerciaux en fichiers numériques d'une manière automatique s'applique à de nombreux processus métier, comme par exemple :

#### Convertir des archives complètes en format PDF/A et PDF interrogeables

ABBYY Recognition Server convertit automatiquement de vastes collections de documents papier, d'images de documents numérisées et de livres complets en fichiers PDF ou PDF/A qui peuvent être ensuite archivés au format électronique, dont la recherche par mot-clé via e-discovery et des systèmes de recherche d'entreprise devient un jeu d'enfant, et que les employés ou les clients peuvent en outre consulter à distance. La solution est très évolutive et peut traiter de grandes quantités de documents à très court terme.

#### Déploiement d'un service de conversion de documents

ABBYY Recognition Server permet de mettre en œuvre un service OCR centralisé au lieu d'installer un logiciel OCR sur différents postes de travail individuels. Tout employé au sein d'une organisation ou d'un groupe de travail peut convertir des documents numérisés vers Microsoft Word ou des fichiers PDF interrogeables, en se connectant directement au service de numérisation ou depuis n'importe quel poste de travail par e-mail ou dossier FTP. Le service peut être déployé au niveau local pour les employés de l'entreprise ou dans un environnement hébergé pour les clients externes.

#### Création de bibliothèques SharePoint interrogeables grâce à la recherche plein texte

Les documents stockés dans Microsoft SharePoint, tels que les TIFF créés par les serveurs de fax et les images PDF créées par les scanners, restent « invisibles » pour le moteur de recherche et ne peuvent être détectés. ABBYY Recognition Server peut récupérer de tels fichiers, les convertir dans un format interrogeable, de type PDF, puis les stocker au même endroit de sorte à ce qu'ils soient intégrés à l'index du moteur de recherche. Si des fichiers PDF numérisés sont enregistrés parmi les fichiers PDF stockés, l'application peut facilement les détecter. Elle ajoute ensuite un nouveau calque de texte, les transformant alors instantanément en PDF interrogeables.

ABBYY Recognition Server peut analyser les bibliothèques SharePoint et les zones de réseau en continu et automatiquement convertir tous les fichiers d'images nouvellement ajoutés. Si une collection TIFF existante doit être conservée dans son format d'origine, ABBYY Recognition Server peut générer du texte interrogeable pour ces images et les transférer sous forme d'un fichier XML vers le moteur de recherche de Microsoft ou Google Search Appliance en laissant les fichiers TIFF originaux à leur emplacement initial.

### FAITS MARQUANTS DU PRODUIT

- Reconnaissance de haute précision des documents dans plus de 190 langues et codes-barres 1D et 2D
- Traitement automatisé de grands volumes de documents dans les délais souhaités
- Copie exacte de la structure du fichier entrant original dans la bibliothèque de documents sortants, avec tous les fichiers au format interrogeable
- Formats d'exportation multiples, dont XML, PDF et PDF/A avec compression MRC élevée, Microsoft Word et autres
- Conversion des documents directement vers Microsoft SharePoint

### BENEFICES

- Résultats OCR fiables grâce aux technologies de reconnaissance d'ABBYY les plus modernes
- Facilité de déploiement avec n'importe quel scanner ou MFP, système de gestion de contenu ou autre système informatique
- Mode de sécurité grâce au dispositif d'équilibrage des charges et la prise en charge des clusters
- Souplesse d'utilisation pour les petites quantités ainsi que pour les volumes importants de documents
- Retour sur investissement rapide grâce au déploiement rapide et la maintenance facilitée

# Conversion automatisée de documents et de PDF

## Aperçu des fonctionnalités clés

ABBYY Recognition Server convertit automatiquement les documents avec une intervention minimale de l'utilisateur. Il opère en arrière plan et effectue de manière autonome toutes les étapes de traitement de documents, 24h/24 ou selon un calendrier prédéfini :

### Étape 1 : numérisation et documents entrants

#### Numérisation

L'application offre une interface de numérisation simple qui prend en charge la numérisation de documents par lots. Elle fournit des outils pour améliorer la qualité des documents comme la prévisualisation et l'amélioration des images, la rédaction manuelle et bien plus encore. Les commandes de script peuvent être utilisées pour, par exemple, scinder automatiquement de grandes pages ou réorganiser des pages suite à une numérisation recto verso.

#### Importation de documents

Les images de documents précédemment numérisées peuvent être récupérées automatiquement à partir de bibliothèques de documents ou reçues par e-mail. Les images de documents importées seront traitées selon les priorités établies et en fonction des ressources informatiques disponibles.



#### Numérisation via TWAIN, WIA, ISIS

Intégration avec tous les scanners en réseau et imprimantes multifonctions (MFP).



#### Surveillance Hot Folder (FTP ou réseau local)

Traite automatiquement les fichiers qui arrivent dans des dossiers prédéfinis.



#### Parcourt les lecteurs réseau et les bibliothèques SharePoint

Détecte les fichiers d'images nouvellement ajoutés et les convertit dans un format interrogeable.



#### Données entrantes par e-mail (Exchange, POP3)

S'intègre avec les serveurs de fax et de messagerie et traite les images en pièce jointe.

### Étape 2 : Traitement des documents

#### Reconnaissance de documents/OCR

Le processus de reconnaissance optique de caractères opère automatiquement sur un poste de travail dédié, la station de traitement. Grâce à la technologie OCR d'ABBYY le système prend en charge un large éventail de fonctions pour accroître la précision de la reconnaissance, dont :

- **Le prétraitement de l'image** (par exemple diviser les doubles pages pour la numérisation de livres ou éliminer le bruit)
- **La définition du type d'impression** (choisir entre le mode texte normal, machine à écrire, matrice à points, OCR-A, OCR-B, MICR E13b)
- **La définition de la langue** (plus de 190 langues et textes historiques avec anciennes polices)

En fonction de la qualité du document et sa structure, le mode de traitement peut être réglé soit sur « précision », soit sur « rapidité ». Pour considérablement augmenter la vitesse de traitement, par exemple pour traiter de nombreux documents dans un très bref délai, on peut ajouter d'autres stations de traitement ou augmenter le nombre de cœurs de processeur.

#### Vérification (facultatif)

Dans certains cas, par exemple lors de la numérisation de livres, la vérification des résultats de reconnaissance s'impose. L'interface intégrée de la station de vérification permet de corriger les résultats, soit sur l'ensemble des documents, soit uniquement sur les documents n'ayant pas atteint un seuil de précision de reconnaissance prédéfini.

#### Indexation (facultatif)

Si nécessaire, l'indexation des documents peut se faire soit manuellement avec l'interface de la station d'indexation automatique ou par un script. Les listes de valeurs d'indices peuvent être importées et synchronisées avec des systèmes tiers.



#### Traitement programmé

Différents types de documents peuvent être traités à différents moments selon un calendrier prédéfini.



#### Mode de sécurité 24/7

Des stations de traitement multiples et le déploiement de clusters peuvent être utilisés pour répartir la charge opérationnelle de manière dynamique et garantir la fiabilité du traitement.



#### Reconnaissance de codes-barres

Permet l'extraction des valeurs de codes à barres 1D et 2D les plus populaires, y compris 2D Aztec, Data Matrix, et les codes à barres de type QR Code.



#### Reconnaissance de textes historiques dans les anciennes polices

Prise en charge de blackletter, Schwabacher et de la plupart des autres polices gothiques en anglais, allemand, français, italien et espagnol

## TRAITEMENT PDF AVANCÉ

- Crée des fichiers PDF et PDF/A avec compression MRC qui réduisent considérablement la taille des documents en couleur.
- Prend en charge le chiffrement : limite l'ouverture et l'impression des documents PDF créés.
- Détecte les PDF numérisés et les fichiers PDF présentant une qualité de texte insuffisante et ajoute un nouveau calque de texte au fichier original.
- Conserve l'image, les signets et les pièces jointes d'origine lors de l'insertion d'un nouveau calque de texte dans le PDF original.
- Les PDF créés numériquement avec un bon calque de texte peuvent être transférés directement vers leur nouvel emplacement.
- Prise en charge des standards d'archivage de documents à long terme : PDF/A-1a, 1b, PDF/A-2a, 2b, 2u
- Crée des PDF optimisés pour le téléchargement Internet.



# Conversion automatisée de documents et de PDF



## Aperçu des fonctionnalités clés – cont.

### Étape 3 : assemblage et export de documents

Après la phase de reconnaissance, ABBYY Recognition Server assemble les pages traitées dans des documents individuels. Les documents peuvent être séparés en utilisant des feuilles blanches ou des pages de code-barres en guise de séparateurs, ou par un nombre fixe de pages par document. La séparation peut alors se faire selon un script précis.

Les documents assemblés dans les formats requis sont transférés vers des emplacements sortants prédéfinis tels que les dossiers réseau, les bibliothèques de documents SharePoint et les adresses e-mail. Mais ils peuvent également être transférés vers d'autres applications connectées via l'API.

En outre, les scripts peuvent être utilisés pour les procédures de routage intelligent et pour transférer les documents à des systèmes de gestion de contenu d'entreprise sur la base des propriétés et attributs des documents. ABBYY Recognition Server prend en charge une variété de formats sortants et permet de créer plusieurs fichiers sortants simultanément.

Pour transformer des archives numériques en archives documentaires électroniques entièrement interrogeables, l'application peut analyser des bibliothèques individuelles, détecter les documents à base d'images non interrogeables et les convertir en format interrogeable. Les documents tels que les courriers Microsoft Word, les présentations PowerPoint® ou les feuilles de calcul Excel qui ne nécessitent pas de traitement peuvent être déplacés vers la bibliothèque sortante à la même position. N'importe quelle bibliothèque de documents peut ainsi être transformée en bibliothèque électronique entièrement interrogeable.



#### Formats sortants multiples

Grande variété de formats, dont PDF et PDF/A interrogeables (compression MRC), XML, RTF, Microsoft Office et autres



#### Publication sur des lecteurs réseau

La structure du dossier d'origine est reproduite automatiquement. Le nom des fichiers sortants peut être défini individuellement, via code-barres, type de document, etc.



#### Envoi par e-mail

Les documents convertis peuvent être renvoyés à l'expéditeur ou à une liste de destinataires spécifiés.



#### Publication sur SharePoint

Les résultats peuvent être téléchargés automatiquement vers des bibliothèques SharePoint. Les PDF numérisés enregistrés dans SharePoint peuvent être complétés d'un calque de texte supplémentaire et enregistrés avec un nouveau numéro de version.

## GESTION ET ADMINISTRATION

### Console d'administration pour une gestion facilitée

ABBYY Recognition Server peut être administré à distance avec la Microsoft Management Console (MMC). Tous les paramètres système, y compris les flux de travail, les licences et les fichiers journaux serveur, sont accessibles de n'importe où.

### Traitement automatisé et programmé

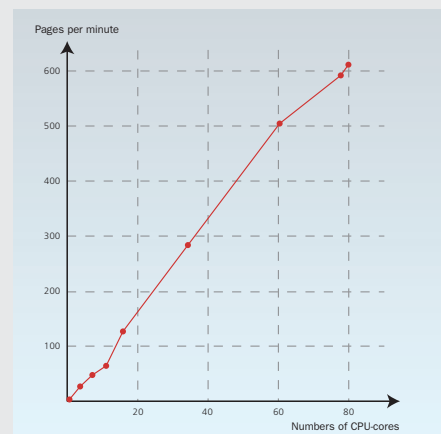
ABBYY Recognition Server traite les documents de manière automatique selon des ensembles de paramètres de traitements prédéfinis (les flux de travail), qui comprennent la source des documents entrants, les étapes de traitement et les paramètres sortants. L'application peut gérer plusieurs flux de travail simultanément selon des priorités établies. Il est possible de programmer l'exécution de flux de travail spécifique à des moments particuliers lorsque la charge opérationnelle est faible (c'est-à-dire dans la nuit).

### Intégration avec les systèmes existants

ABBYY Recognition Server peut être facilement connecté à des applications externes, comme par exemple les archives numériques de systèmes de gestion de contenu d'entreprise, via des tickets XML, des API basées sur la technologie COM et API Web Service.

### Evolutivité et flexibilité

Il est possible d'utiliser un ordinateur multicœurs ultraperformant en guise de station de traitement pour accroître la vitesse de traitement, ou de distribuer la charge opérationnelle sur différents ordinateurs au sein du réseau. L'architecture évolutive et flexible permet la mise en œuvre de systèmes pouvant facilement traiter des centaines de pages par minute.



Le fait de déployer des performances de traitement supérieures au niveau du processeur permettra d'accroître la vitesse de traitement.

Ces informations s'appuient sur des tests internes. Les performances système peuvent varier en fonction de la qualité des images, des performances du matériel, de la configuration du réseau et d'autres facteurs.

## SPÉCIFICATIONS

### Configuration système requise :

- **PC** avec Intel® Core™2/2 Quad/Pentium® / Celeron®/Xeon™, AMD K6/ Turion™/ Athlon™/ Duron™/ Sempron™ processeur cadencé à une vitesse d'horloge minimale de 2 GHz
- **Système d'exploitation** : Microsoft® Windows® 8, Windows 7, Windows Vista®, Windows Server® 2012, 2008
- **Mémoire** :
  - Gestionnaire de serveur : 1 Go de RAM
  - Station de numérisation : 1 Go de RAM
  - Station de traitement : 512 Mo de RAM et 300 Mo supplémentaires pour chaque processus de reconnaissance
  - Station d'indexation : 768 Mo de RAM
  - Station de vérification : 1024 Mo de RAM
- **Espace disque dur** :
  - Gestionnaire de serveur : 20 Mo pour l'installation, 1 Go supplémentaire pour l'utilisation du programme
  - Station de numérisation : 1 Go
  - Station de traitement : 600 Mo pour l'installation et 1 Go pour l'utilisation du programme
  - Station d'indexation : 500 Mo pour l'installation et 1 Go pour l'utilisation du programme
  - Station de vérification : 700 Mo pour l'installation et 700 Mo pour l'utilisation du programmeUn espace disque supérieur peut être nécessaire pour utiliser le programme, en fonction de la complexité, de la qualité et du nombre d'images. La configuration requise varie selon les composants serveur et les modules supplémentaires utilisés. Contactez ABBYY pour plus de détails sur les configurations.
- Microsoft .NET Framework 3.5 ou version ultérieure pour enregistrer les fichiers sur Microsoft SharePoint Server
- Microsoft Outlook 2000 ou version ultérieure, pour traiter les messages électroniques.
- Microsoft IIS 5.1 ou version ultérieure (pour API Web)
- Scanner prenant en charge TWAIN, WIA ou ISIS

### Langues d'interface utilisateur\*

Anglais, français, allemand, italien, espagnol, russe, portugais (brésilien), tchèque, hongrois, polonais, chinois (traditionnel et simplifié)

\* La release 1 ne propose l'interface utilisateur qu'en russe et anglais.

### Formats d'entrée

- BMP, PCX, DCX, GIF, TIFF / TIFF multipage, WDP, WMP
- JPEG, JPEG 2000, JBIG2, PNG, RLE
- PDF (jusqu'à version 1.7), DjVu, JPX

### Langues d'OCR

Plus de 190 langues

### Types d'impression

Normale, Fax (mode pour texte à faible résolution), machine à écrire, imprimante matricielle, OCR-A, OCR-B, MICR (E13B), gothique

### Types de codes-barres

- 1D** : Check Code 39, Check Interleaved 25, Code128, Code 39, EAN 13, EAN 8, Interleaved 25, CODABAR (sans somme de contrôle), UCC Code 128, Code 2 of 5 (Industrial, IATA, Matrix), Code 93, UPC-A, UPC-E, Patch Code et Postnet
- 2D** : PDF 417, Aztec, Data Matrix, QR Code

### Formats de sortie

#### Formats éditables

- RTF, TXT, HTML, CSV, EPUB
- DOC, DOCX, XLS, XLSX
- XML, Alto XML, FineReader dans le format interne

#### Formats interrogeables

PDF (jusqu'à version 1.7); PDF/A

#### Formats d'image

- PDF image seulement PDF, PNG, JBIG2
- JPEG, JPEG 2000, TIFF

### Options d'intégration et de personnalisation

XML Tickets, API basées sur la technologie COM et API Web Service, scripts en VBScript et JScript

## ABBYY Recognition Server: Editions du produit

### Edition étendue

L'édition étendue offre toute une palette de fonctionnalités OCR, des possibilités d'intégration avec les applications externes et d'implémentation dans le cadre d'une architecture de services Web. Cette édition propose également des langues OCR supplémentaires (thaï et hébreu), l'exportation vers le format XML, l'API Web Service, les API basées sur la technologie COM ainsi que la prise en charge de l'intégration des XML Tickets et de Microsoft SharePoint. Sur demande, la reconnaissance des lettres gothiques dans les documents historiques (police Black letter) et l'OCR pour l'arabe, le chinois, le japonais et le coréen sont également disponibles.

### Edition professionnelle

L'édition professionnelle offre des fonctionnalités standard pour les organisations qui souhaitent un service OCR en tâche de fond automatisé au déploiement simple et ne souhaitant pas d'intégration au sein d'autres applications. Cette fonctionnalité peut être étendue en toute flexibilité grâce à un ensemble de modules complémentaires.

**Remarque :** Veuillez consulter la liste tarifaire ABBYY Recognition Server pour connaître le détail des prix, les fonctionnalités disponibles et les possibilités d'extension des éditions professionnelle et étendue.

## Licenses

Les deux éditions sont disponibles dans le cadre d'un système modulable de licence. La licence définit l'ensemble des fonctionnalités qui peuvent être étendues de manière modulable sans réinstallation logicielle. ABBYY Recognition Server permet donc à ses clients d'accroître l'efficacité de leur processus, qu'il s'agisse de faible volume de traitement de documents ou du traitement de millions de pages pour les grandes entreprises ou les prestataires de service de numérisation.

### Licence par processeur

La plupart des utilisateurs ont besoin d'une licence par processeur. Le nombre de cœurs de processeur influence la vitesse de traitement : plus il y en a, plus le processus de conversion sera rapide. Au besoin, les utilisateurs peuvent à tout moment ajouter des processeurs supplémentaires.

### Licence par nombre total de pages

Pour les projets à caractère unique (c'est-à-dire la conversion des archives d'une société au format PDF/A), nous recommandons de faire l'acquisition d'une licence par nombre total de pages, qui correspond au nombre de pages devant être traitées. Le nombre de cœurs de processeur n'étant pas limité, même les archives documentaires très importantes peuvent être converties très rapidement.

## Modules complémentaires

La fonctionnalité peut être étendue en ajoutant l'option numérisation, indexation et station de vérification ou l'assistance de cœurs de processeur supplémentaires. Possibilités d'étendre la licence à d'autres langues de reconnaissance pour accroître les fonctionnalités. Possibilité d'ajouter des connecteurs pour Google Search Appliance et pour Microsoft Search pour autoriser ce système à détecter des images grâce à la recherche plein texte.

## Maintenance logicielle

La maintenance logicielle protège l'investissement logiciel. Grâce à la maintenance logicielle, le client bénéficie de mises à jour et de mises à niveau logicielles régulières pour profiter des toutes dernières versions ainsi que d'une assistance technique gratuite. La maintenance logicielle correspond à 20 % du prix de la licence et est facturée séparément sur une base annuelle. La maintenance logicielle est obligatoire.

## Versions d'évaluation

ABBYY propose des versions d'évaluation entièrement opérationnelles. Veuillez contacter ABBYY ou ses partenaires pour de plus en plus d'informations ou demander votre version d'évaluation sur [www.france.abbyy.com/recognition-server](http://www.france.abbyy.com/recognition-server)



**ABBYY Europe GmbH**  
Elsenheimerstr. 49,  
80687 Munich, Allemagne  
Tel: +49 89 511 159 0  
sales\_eu@abbyy.com  
[www.ABBYY.com](http://www.ABBYY.com)

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

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

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

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

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

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