



ABBYY Recognition Server™ for the Google Search Appliance™



The ABBYY Recognition Server™ for the Google Search Appliance™ (GSA) works as a background optical character recognition (OCR) service that helps organizations unlock and access documents saved in image-based formats. ABBYY's Recognition Server for GSA gives enterprise organizations an automatic and cost-effective way to organize and access knowledge that was hidden from search results in the past.

The Need

Immediate Information Access Speeds Productivity

In order for organizations to more efficiently share knowledge and increase productivity, access to information from its archive of documents needs to be readily available—regardless of storage system, segmented archives or various databases. This type of unrestricted and immediate access increases employee productivity and empowers knowledge workers creating a more flexible, mobile, and efficient organization better able to seize new market opportunities.

The Technology

ESS Enables Unstructured Document Searches

Enterprise Search System (ESS) Technology provides an organization's knowledge workers with a powerful means of searching valuable electronic content from a unified search box, increasing their effectiveness even if they don't know where to look. The three components of ESS—crawling, indexing, and search engines—work well for documents that exist in a format immediately suitable for full text indexing.

Google's ESS: Google Search Appliance

Google developed its own ESS designed to index searchable and editable file formats on an organization's file servers, Web servers, content management systems, and other resources. The intuitive user interface of the GSA simplifies the search for various formats of documents or information within an organization.

The OCR Advantage

More Thorough Searches with More Formats

The OCR technology provided by ABBYY Recognition Server for the GSA adds more document formats to the inventory of file types handled by GSA today. It extends GSA's indexing capabilities beyond popular formats such as HTML, DOC, PDF, XLS, and TXT, to support documents in TIFF, JPEG, and other image formats to make them available for search.

ABBYY's leading OCR accuracy builds on the GSA's industry-leading relevancy to ensure that the GSA search index is of the highest quality, empowering accurate queries and generating confidence in the results. Global organizations generate and search for content world-wide and in multiple languages, therefore OCR supporting enterprise solutions must be flexible and multi-lingual. ABBYY's OCR supports over 191 languages, including Chinese, Japanese, and Korean, making it the only choice for complex and global OCR and ESS initiatives

Key Features and Benefits

No document goes undetected

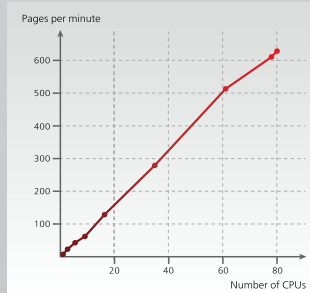
GSA together with Recognition Server crawls document libraries and indexes the content of all stored documents—both text-based and image files—so no documents go undiscovered.

The right document, right away

Award-winning ABBYY OCR technology in Recognition Server will ensure that text content is extracted precisely even from low-quality images.

Unmatched speed & scalability

Recognition server is installed on separate servers on the network, taking the resource-intensive OCR task off the GSA, to speed the indexing process. Additional servers can be added to the system to meet any volume needs.



Powerful OCR server at your service

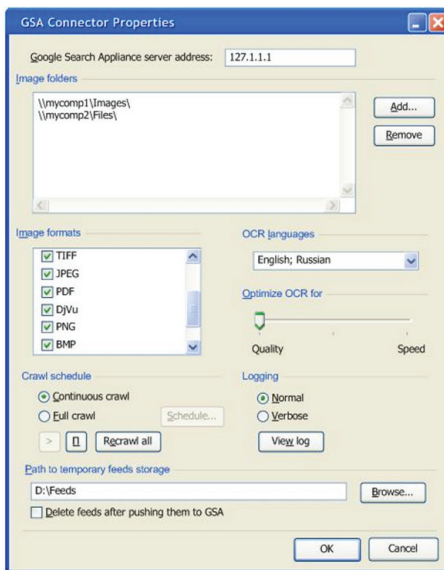
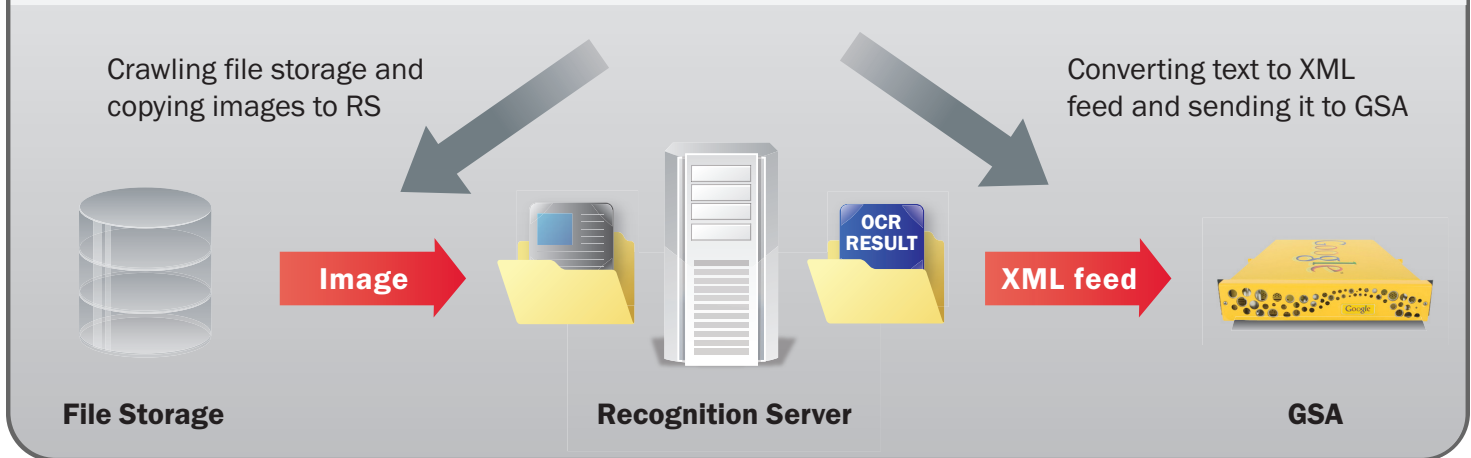
Purchasing Recognition Server with a connector to GSA also allows its full use for any other OCR or document conversion task, either within a company's business process or ad hoc.

Multilingual searches

ABBYY is the leader in linguistic recognition. Recognition Server supports over 191 languages, making it the clear choice for global search initiatives.



Connector



How it works

ABBYY's Recognition Server for the GSA runs automated and unattended, extracting text results from image-based documents and delivering results to the GSA. The end result: More productive and relevant searches without limits on document files formats. The system searches specified network file shares, including nested subfolders, picks any image files and sends them to the Recognition Server for OCR. It leaves the electronically created documents, such as DOC or XLS, for indexing by GSA. ABBYY Recognition Server for the GSA also picks up the OCR results and submits the recognized text of the document to GSA in an XML feed, accompanied with a link to the original image.

ABBYY's Recognition Server for the GSA allows the following configurations:

- Control the selection of image types to be OCR'd
- Assign the specific file shares to be monitored
- OCR optimization for up to 191 languages
- Define crawling and processing schedules
- Log monitoring settings, and more...

System Requirements

• Google Search Appliance

• Recognition Server for the GSA Processing Station

- PC with 500 MHz or higher CPU
- Operating System: Microsoft® Windows® Server 2008, Windows Vista®, Windows 7®, Windows Server 2003, Windows XP, Windows 2000
- Memory: 128 MB RAM plus an additional 100 MB for each recognition process
- Hard disk space: 230 MB for installation and 100 MB for program operation*

* **Note:** The hard disk space required for program operation may be larger depending on the complexity, quality and number of the images.

• Recognition Server for the GSA Server Manager

- PC with 500 MHz or higher CPU
- Operating System: Microsoft® Windows® Server 2008, Windows Vista®, Windows 7®, Windows Server 2003, Windows XP, Windows 2000
- The account which is used by Recognition Server must have read/write permissions to the following registry branches:
 - HKEY_CLASSES_ROOT
 - HKEY_LOCAL_MACHINE\Software\ABBYY
 - HKEY_CURRENT_USER\Software\ABBYY
- Memory: 128 MB RAM
- Hard disk space: 15 MB for installation and 100 MB for program operation*

http://www.abbey.com/recognition_server/gsa/

© 2009 ABBYY. All rights reserved. ABBYY, the ABBYY Logo, FineReader, Lingvo, Recognition Server and FlexiCapture are either registered trademarks or trademarks of ABBYY Software Ltd. Other product names mentioned herein may be trademarks and/or registered trademarks of their respective owners and are hereby recognized. All other trademarks are the property of their respective owners. Information contained in this document is subject to change without notice.

ABBYY USA
880 North McCarthy Blvd.,
Suite #220
Milpitas, CA 95035, USA
Tel +1 866.463.7689
Fax +1 408.457.9778
sales@abbyyusa.com

www.ABBYYUSA.com