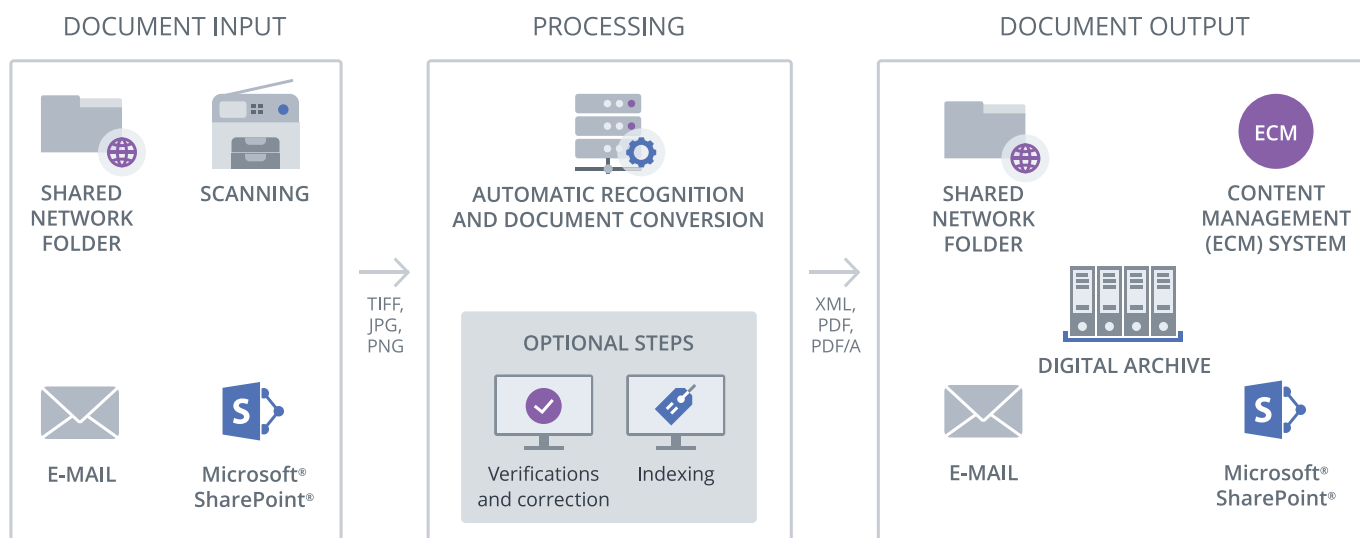# FineReader®
Server

Converting documents into accessible, searchable, sharable formats — such as PDF®, HTML, and XML — lets you store information electronically across your IT system, enabling greater efficiency through faster information retrieval.

Designed for high-volume document conversion, FineReader Server automatically converts large collections of documents into searchable, sharable digital libraries. Our server-based OCR and PDF conversion offering converts scanned and electronic documents into PDF, PDF/A, Microsoft® Word, or other formats for search, long-term retention, collaboration, or additional processing — quickly, accurately, and automatically.

FineReader Server receives document images from shared network folders, scanners, e-mails, and Microsoft® SharePoint® and automatically converts them into searchable digital formats using optical character recognition (OCR) technology. If needed, the user can can manually correct text information and/or add metadata to the document. The resulting digitized files can be saved to any number of storage areas and/or delivered to other applications.

Benefits overview

FineReader Server converts documents automatically, with minimum user intervention. It runs in the background and performs all document processing steps independently — around the clock or at pre-defined times.

Standardize your content
Transform collections of documents into standardized, well-organized digital libraries.

Realize rapid ROI
FineReader Server is fast to deploy and easy to maintain, so you can start seeing bottom-line results faster.

Schedule processing
Enable conversions as needed around the clock or schedule batch processing to optimize use of hardware resources.

Empower non-technical business users
FineReader Server requires no special training or previous knowledge to start conversion processes.

Streamline document handling
Create digital documents that can be easily stored, quickly distributed to specialized systems, or shared among teams.

Make content accessible, searchable
Business users can quickly search digital archives for documents containing relevant keywords.

Key features

Trusted character recognition
Delivers fast, accurate results in over 190 languages.

High-volume scalability
Converts large volumes of documents in a short time frame.

Server-based architecture
Uses all available hardware resources in the most efficient way possible.

Flexible PDF technology
Compresses PDFs to minimize file size while preserving quality; supports PDF/A-1a, PDF/A-1b, PDF/A-2u, and PDF/A-3a formats.

Wide range of supported formats
Automatically converts from PDF, JPEG, TIFF, Word, Excel, OpenDocument Text, PowerPoint®, HTML, and other formats

Integration with SharePoint
Automatically converts documents from SharePoint libraries into searchable PDFs.

Barcode recognition
Reads 1D and 2D barcodes to enable document separation and/or addition of metadata.

Document types and metadata
Auto-assigns document types and attributions; allows for manual creation of metadata if needed.

Recognition of historical fonts
Supports black letter, Schwabacher, and most other Gothic fonts in English, German, French, Italian, and Spanish.

Integration into existing systems
Connects easily with digital archives or enterprise content management systems via XML tickets, COM-based APIs, and web service APIs.

Document separation
Automatically separates documents based on number of pages, blank pages, barcode pages, and/or scripted rules.

Flexible licensing and pricing
Offers licensing models for any size organization or project, with flexibility to adjust as needs evolve.

How it works

1
Step

Document input

FineReader Server receives document inputs from scanners, shared network folders, emails, and Microsoft SharePoint.

Scanning

FineReader Server offers an easy-to-use Scanning Station interface that supports scanning of batched documents. Onboard quality improvement tools include image preview and enhancement, manual redaction, and others. Users can leverage scripting commands, for example, to auto-split large pages or re-order pages after duplex scanning.

Document import

FineReader Server can automatically retrieve previously scanned images from document libraries and files sent as e-mail attachments. Imported document images will be processed with corresponding priorities and according to available computing resources.

- Scanning via TWAIN, WIA, ISIS
- Integrates with all network scanners and MFPs
- Hot folder watching (FTP or local network)
- Automatically processes files arriving in defined folders
- Crawling of network shares and SharePoint libraries
- Detects newly added files and converts into searchable formats
- Input via e-mail (Microsoft Exchange, POP3, IMAP)
- Integrates with fax and e-mail servers and processes image attachments

Multiple input formats:

Image files

- TIFF / Multipage TIFF
- Compression methods: Uncompressed, CCITT3, CCITT3FAX, CCITT4, PackBits, JPEG, ZIP, LZW (8/24bits)
- JPEG, JPEG 2000
- JBIG2
- BMP
- GIF
- PNG
- WDP
- XPS
- PCX, DCX

Office documents

- PDF
- DjVu
- DOC, DOCX, ODT
- XLS, XLSX, ODS
- PPT, PPTX, ODP
- TXT, HTML, HTM, RTF

Emails

- Mail servers supported protocols: IMAP, MAPI, POP3, like MS Exchange, Google® Mail, IBM® Domino, etc
- ail message files stored in file system, like MSG and EML files

Customization

Ability to create plugins for expanding list of supported formats, e.g. convert CAD file to PDF with help of Autocad® or other software

2 Step

Document processing

FineReader Server processes document images via automatic recognition and document conversion, with optional verification and indexing capabilities.

Document recognition/OCR

FineReader Server's OCR process runs automatically on a dedicated workstation, the Processing Station. Using ABBYY's highly accurate OCR technology, FineReader Server supports a broad range of functions to increase recognition accuracy, including

- Image pre-processing (for example, splitting dual pages for book scans or clearing background noise)
- Print type definition (normal text, typewriter, dot-matrix, OCR-A, OCR-B, MICR E13b, and Gothic)
- Language definition (auto-recognition of more than 190 languages and historic texts in old fonts)

Depending on the document's quality and structure, processing mode can be set on either "precision" or "speed." To increase processing speed significantly — for example, to process many documents within a tight time frame — additional Processing Stations or a greater number of CPU cores can be added.

FineReader Server's scheduled processing feature enables it to process different kinds of documents at different times according to a predefined schedule.

Verification (optional)

In some cases — for example, when digitizing books — recognition results must be verified. FineReader Server's integrated Verification Station interface offers the option of correcting results, either on all documents or only on documents that did not reach a pre-defined recognition accuracy threshold.

Indexing (optional)

If required, document indexing can be done either manually — using the Indexing Station interface — or automatically using a script. Lists of index field values can be imported and synchronized with third-party systems.

3 Step

Document assembly and export

FineReader Server assembles processed pages into individual documents. Documents can be separated three different ways:

- Using blank sheets or barcode pages as separators
- According to a fixed number of pages per document
- According to a scripted rule

Assembled documents in the required formats are delivered to pre-defined output locations — such as network folders, SharePoint document libraries, and e-mail addresses — or passed to other applications connected via API.

Scripts can also be applied for intelligent routing and delivery of documents to enterprise content management systems based on document properties. FineReader Server supports a variety of output formats and enables creation of several output files at the same time.

FineReader Server can crawl individual libraries, detect non-searchable image-based documents, and convert them into searchable formats. Documents such as Microsoft Word files, PowerPoint presentations, or Excel spreadsheets, which require no processing, can be moved into the output library to the same position.

Multiple output formats:

- PDF, PDF/A-1a, PDF/A-1b, PDF/-2a, PDF/A-2b, PDF/A-2u, PDF/-3a, PDF/A-3b, PDF/A-3u, PDF/E, PDF/UA
- RTF
- DOC, DOCX
- XLS, XLSX
- TXT, CSV
- HTML
- TIFF
- JPEG, JPEG 2000
- PNG
- EPUB
- XML
- Alto XML
- FineReader internal format (FineReader Engine-compatible)