



Adlib ExpressRecognition Server

Convert image files into searchable PDF documents

Extract maximum value from your paper documents

Adlib ExpressRecognition Server provides Optical Character Recognition (OCR) from a centralized server. It quickly converts millions of static electronic documents such as faxes, photocopies and images into searchable electronic files such as Text and PDF. And because this is an Express Server product, you can be sure that it will integrate into almost any automated document workflow and there are no limits or additional costs associated with the number of documents, pages or users.

Workflow integration options

As a server-based application ExpressRecognition Server is a service accessible to multiple workflows and applications. Adlib Software designed all of our products with multiple integration options.

Highlights

- Industry leading accuracy
- 119 Languages Supported
- Customizable user dictionaries
- Integration with Express' powerful Conversion and Publishing tools
- Zonal OCR
- 1d and 2d Barcode Recognition
- Optical Mark Recognition

| | |
|-----------------|--|
| XML Job Tickets | A flexible set of instructions that tells Recognition Server what to do, how to do it and any other detailed instructions that can possibly be executed by the server. |
| Watched Folders | Configure single or multiple, recursive folders, each with custom processing configurations that can be shared by multiple applications and workflows |
| Email | Adlib ExpressRecognition Server is capable of receiving image documents by email and returning the searchable results to the sender as an email attachment. |
| Web Services | SOAP compliant messaging makes it possible to easily connect to Recognition Server through industry compliant web service interfaces. |

These flexible integration options are simple to set up and easy to use. Adlib Software also provides the tools and the experience to guide you through the process. For example, once XML Job Tickets are built, it is easy to create a library of these “templates” to be called upon by other applications or by other unique Job Tickets.

Why server-based OCR

Companies use Adlib ExpressRecognition Server to archive image-only document files and capture knowledge in non-searchable documents. By extracting the text from an image and adding it to that file, that document becomes fully searchable, making it much more accessible within an archive. Image only documents such as legacy content can be more readily repurposed by mining the text and applying it to other documents. Most importantly, by having the OCR capabilities available as a service, manual intervention is dramatically reduced – saving time, effort and the high probability of costly manual errors. Adding Relay, a browser-based application and UI front end to Express Server applications, on-demand OCR can be performed without affecting desktop performance because all the processing is done on the server.

Reliability and Scalability

As part of the Adlib Express Server line, Recognition Server has built-in fault tolerance and error handling. The Fault Monitoring and Recovery component, which is found in all Adlib Express Server products, is a configurable module that monitors the server for interruptions and delays. If desired, the system can restart itself, retry the job, place the corrupted file in a reserved folder and generate an error log or any combination of the above. As an extension of this functionality, if more than one server is registered on the network, Express Server will automatically reroute unfinished jobs to other servers in an effort to ensure job completeness. In a similar manner, multiple jobs can be routed to multiple servers to take advantage of economies of scale and automatically speed up the overall process.

Speed and accuracy

There are many desktop OCR applications on the market today but very few server-based ones. Taking advantage of the benefits offered through server technologies, ExpressRecognition Server offers superior performance and accuracy. Built-in features for deskewing and despeckling images result in even higher levels of accuracy. Adjustments can be made to balance speed with accuracy when necessary.

Integration with other Adlib Express Server products

Adlib ExpressRecognition Server is an integral component of the Adlib Express Server line of Document Transformation Technologies. Implemented as part of the overall solution, it is an invaluable tool in the document conversion, recognition and publishing framework. Combined, these interoperable solutions make it possible to automatically convert hundreds of document file types, including image only files, into PDF, TXT, XPS, RTF, WORD, JPG and TIFF files, convert images to searchable text and publish results with headers, footers, watermarks, page numbers, tables of contents and bookmarks. Multiple documents can be assembled into one document or a single document can be split up into predefined volumes with slip sheets and tabs. All this functionality can be achieved through one solution that integrates seamlessly into your document workflows.

Adlib ExpressRecognition Server Highlights:

| | |
|--|---|
| Broad file type support | Wide support for different image types such as TIFF and JPG in addition to PDFs ensures that maximum value is extracted from image based documents. |
| Multi-lingual recognition | Recognition Server recognizes 119 Languages using 17 Language Dictionaries to ensure the highest possible accuracy. |
| Document processing can be optimized for accuracy or speed | The recognition process can be optimized for speed or accuracy to support the requirements for all document workflows. |
| Image pre-processing functions include de-skewing and de-speckling | De-skew and De-speckle are settings that can be used to optimize the accuracy of the recognition process. De-skew automatically straightens the document while de-speckle removes black dots from the original image. |
| Advanced error handling | Documents that exceed a time or error threshold can be identified and optionally routed to an error folder. |
| Data extraction | Adlib ExpressRecognition Server can be configured to output recognized text in both XML and CSV formats. |