Intelligent OCR. Highest Accuracy.

Increase efficiency by automating creation of searchable pdf content

Faxes, photocopies, TIFFs and other scanned images continue to represent a significant portion of enterprise content. The challenge is that even when digitized these are completely unsearchable, rendering their content useless to other applications and analytics software.

Adlib's intelligent OCR (Optical Character Recognition) automates the rendering of image-based content into searchable PDF files within any document workflow. Organizations can leverage the Adlib OCR to recognize image-only document files creating searchable PDF documents suited for downstream processes including long-term archival, classification, migration and more.

By accurately deciphering text characters from images and creating PDF files, content becomes fully searchable, enabling easier access for every user in a workflow process where the information is required.

When performance, document fidelity and flexibility are important, the world's largest and most successful companies choose Adlib.

Intelligent OCR as a Centralized Shared Service

By making intelligent OCR capabilities available as a centralized service, organizations effectively mitigate the risks associated with manual error, while saving significant time and money. When performance, document fidelity and flexibility are important, the world's largest companies choose Adlib.

Why Adlib?

FUNCTIONALITY	ADLIB OCR	OTHER OCR
Fidelity	Highest definition in the industry—bookmarks, annotations, hyperlinks, watermarks, logos, and images retained	Definition low, not able to add or retain enhancements such as watermarks
Compliance	Does not just convert source document into a new document. We expertly create the text layer and intelligently apply it to the source document leaving it completely searchable and compliant to document archive organizations such as the US National Archives and Records Administration (NARA)	Scans paper to flat file with no hyperlinks, bookmarks, or metadata
Page Cleanup	Automated page cleanup (pages are de-skewed and/or rotated to optimize recognition accuracy)	Limited as many scanners take an image of the page and are unable to de-skew and/ or rotate the page to optimize recognition
Archive	Can add long term digital preservation to workflow by automatically converting to any version of PDF/A	Often requiring operator to manually enter metadata, and other information
Formats	300+ formats including legacy and specialty (e.g. CAD) formats	Limited file format support
Large Page Size	Can OCR large page sizes in any format	Don't support large page sizes
Multi-Lingual	Recognizes over 115 languages, including Asian languages	Varies but most have extremely limited lanuage support
Language Dictionaries	Can perform word correction for 17 languages	Don't support word correction
Vertical Dictionaries	Can perform word correction for Legal, Financial and Medical terms for 4 languages	Don't support word correction for vertical terms
User Defined Dictionaries	Supports adding dictionaries for word correction for words used in your organization in 112 languages	Don't support User Dictionaries
CAD	CAD files can be OCR'd including very large pages	Unable to handle CAD or large pages
Metadata	Automatically extracts advanced intelligent metadata, not just title and author	Require manual metadata entry
Content Extraction	Configured the recognized text to feed automated workflows including content flow identification for analysis and reporting activities	Manual intervention necessary to extract content
Compression	Supports JBIG2 and Mixed Raster Compression (MRC) automatically applies the best compression to minimize PDF file size for optimum document transport	Limited compression support and no compression optimization
Redaction	Most powerful PDF document redaction in the industry to removes sensitive content	Most are unable to handle redaction but those that do may only mask as opposed to permanently remove the content
Digital Documents	Digital documents can be captured without transforming born digital document to raster format such as TIFF and requiring OCR and converted to high definition (HD) Adlib PDF	Digital conversion to raster image resulting in low quality image and searchable content
Error Handling	Advanced error handing (e.g. complex or large documents that exceed a time or error threshold are identified and routed to appropriate processes)	Manual error handling
Multi Repositories	Adlib has connectors that provide integration into the top 5 repositories and can provide integration into any of the available repositories	Many are only integrated with a single repository
Load Balancing	Supports load balancing and high availability to provide 24/7 service and meet Service Level Agreements (SLAs)	Many do not support job monitoring or load lalancing