

### SUMMARY

OnBase OCR (Optical Character Recognition) is a highly accurate full-page character recognition package working with the OnBase Document Imaging or Document Import Processor modules. OnBase OCR creates readable text renditions of image document, to facilitate text searches. Searchable PDF or MS Word formatted renditions can also be created. Users are then able to instantly locate specific words or phrases within a document or can locate a specific document from within a group of documents.

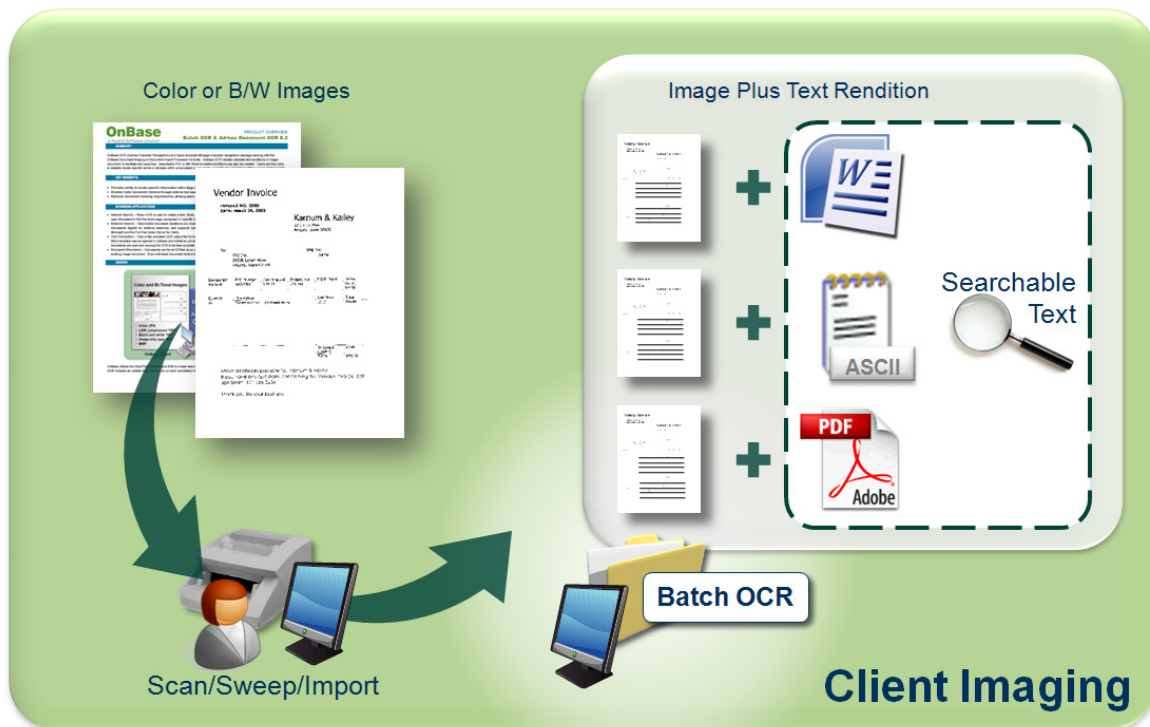
### KEY BENEFITS

- **Provides ability to locate specific information** within large image documents.
- **Enables faster document retrieval** through external text search across many documents.
- **Reduces document indexing requirements**, allowing users to search document content directly.

### BUSINESS APPLICATIONS

- **Internal Search** – When OCR is used to create a text, Word, or PDF rendition of a document, user can perform searches on the open document to find the exact page, paragraph or specific data that they need.
- **External Search** – Searchable document renditions are created from scanned batches of image documents. This makes these documents eligible for external searches, and supports optional full-text search engines like the Full-Text Index Server for Microsoft and the Full-Text Index Server for Verity.
- **Text Correction** – One of the available OCR output file formats is Microsoft Word. When this output format is selected, the MS Word rendition can be opened in OnBase and edited so corrections can be made. This especially useful when old or poor quality documents are scanned causing the OCR to be less accurate.
- **Document Revisions** – Documents can be re-OCR'd as an ad hoc task, especially when pages have been scanned to an existing image document. Even individual document revisions can have separate and distinct OCR renditions.

### DESIGN



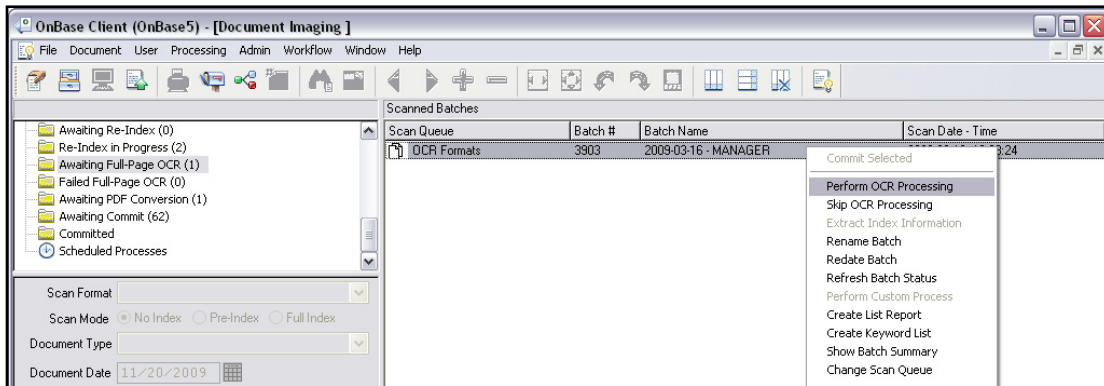
OnBase utilizes the OmniPage OCR engine SDK to create searchable document text renditions in selectable file formats. OnBase OCR includes an installer that must be run on each workstation that will perform OCR.

## KEY FEATURES

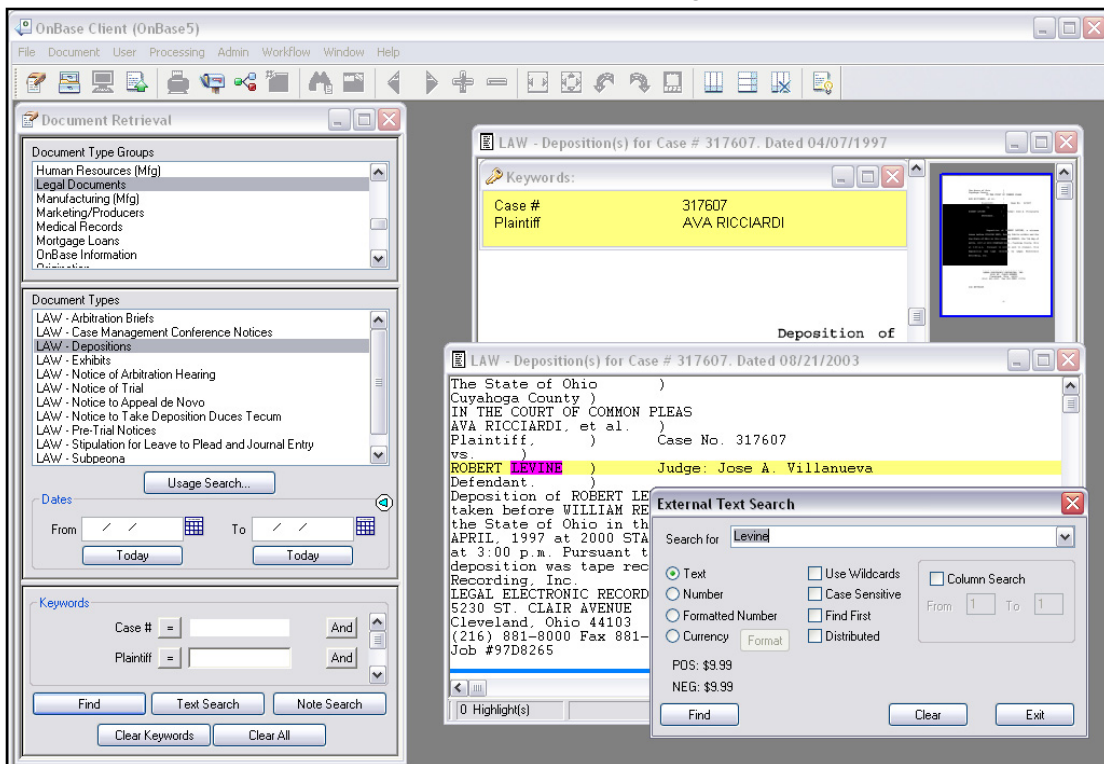
- **Selection of output formats** for the OCR renditions, including ASCII Text, Word, and searchable PDF.
- **Supports OCR in 18 languages**, all from a single workstation.
- **Supports OCR on color and bi-tonal images.** JPG, JPG compressed TIFF images, and bi-tonal TIFF images are supported. Image-only type PDF files can also be converted to searchable PDF format.
- **Documents can be sent for OCR**, from the import dialog or from a right-click option.
- **Ad-Hoc Document OCR** is available for specific users that need to occasionally OCR individual documents or re-OCR documents. This can be performed from a document hit list on any supported image format.

## INTERFACE

### Batch OCR on a scanned batch



### Searchable text rendition for an image document



**HYLAND**  
**SOFTWARE**