ImageRamp 8.x
Search and Index

DocuFi
www.DocuFi.com

Revision Date Sept. 2021

Contents

Introduction

General Information

ImageRamp Indexer Software is a search and retrieval solution for documents that are scanned or existing text format based.

Supported File Formats

- **Input:** Monochrome (black and white), color or greyscale TIF files; most Adobe® Acrobat files (black and white, color, or greyscale PDF and text, image, and OCR PDF files); color or greyscale JPG and PNG files; and MS-Word .docx files are supported as input files. .

 ***CAUTION:*** Please test all file types in your environment.

System Requirements

The ImageRamp Batch family supports viewing and manipulation with:

- Windows 10/8/7, and Windows 2008/2012 Server and newer platforms
- Microsoft .NET Framework 4.6 and above
- 150 MB of free disk space

Conventions Used in this Documentation

These are the conventions and type-styles used in this manual:

- Items presented in ***Italics***:
  Screen, field, checkbox, menu and tile names.
  Terms and concepts when first introduced in the text.
  Variables, file names and numeric values.
  Special notes and important messages.
- **[RETURN]** indicates the action of pressing the 'Enter' or 'Return' key.
- Keyboard entries to be typed and buttons/icons to be clicked are printed in lower case **Arial bold font**.
- Clicking refers to a left mouse click unless otherwise noted as a "right-mouse click"

Trial Version Limitations

A full version of ImageRamp Batch is valid for 30 days.

Getting Help

Need help with ImageRamp? There are number of ways to get the help you need.

Help and Documentation - This help documentation includes details on commands, features and functions of the software.

DocuFi Website – The DocuFi website has ordering information and more. Visit http://www.docufi.com.

Getting Help by E-Mail - Send your questions or comments to us: support@docufi.com. A Customer Service Representative can provide technical support or answer your inquiry by e-mail.

Installation, Opening/Closing the Application and Help

Installing ImageRamp Indexer

Install the application:

1. Launch the ImageRamp installer file. This program will install ImageRamp Batch and support components into the "Program files (x86)\ImageRamp8" folder. It will also install the ImageRamp Scan Separator which creates barcode separation pages and ImageRamp Composer. These may be launched from the desktop icons.
2. Follow the standard prompts and instructions.

This will install the ImageRamp Batch product family which will function in trial mode until it is licensed. To fully activate the program it must be authorized with a license key.

Opening and Closing the Application

Start the application with the icon placed on the desktop or by selecting the program through Windows program menu. To exit the application, select the **Exit** icon in the upper right corner of the application screen.

Authorizing ImageRamp Indexer

To activate the product once it has been installed, the license must be authorized with a license key. Email or call DocuFi if you have not automatically received your key in an email after purchase. To access the application in trial mode, simply click on the **Request a License Key** button. You will receive an email with a trial key that can be entered and activated with the **Activate Trial/License** button.

Illustration: ImageRamp Registration Screen displayed on installation.

NOTE: The trial version of ImageRamp Indexer is valid for 30 days. Contact DocuFi for a full product online demonstration.

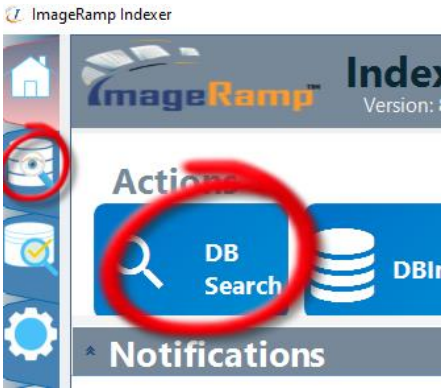
Accessing Application Help

To access the application help, select the ***Help Manual*** tile from the ***Learning Center*** panel of the ***Home Tile*** screen.

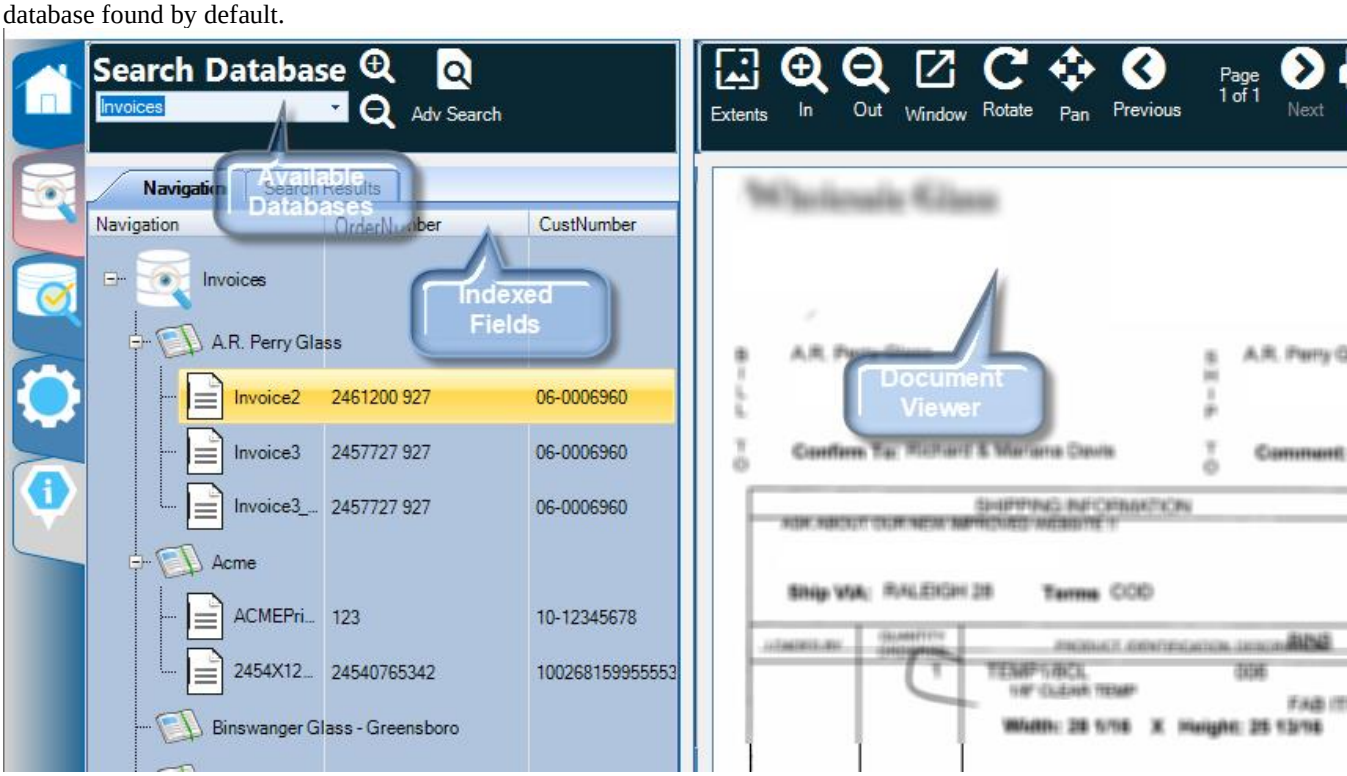
Database Searching

Searching for Documents already Indexed

ImageRamp provides a simple and intuitive way to search for documents indexed into a SQL Server or Access database. To i
the Search icons from the home page or the left menu items as shown below.

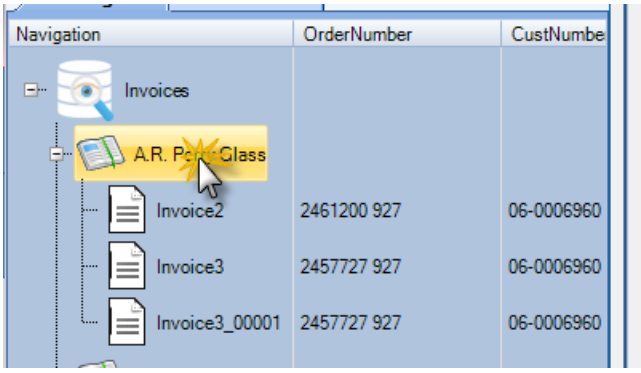


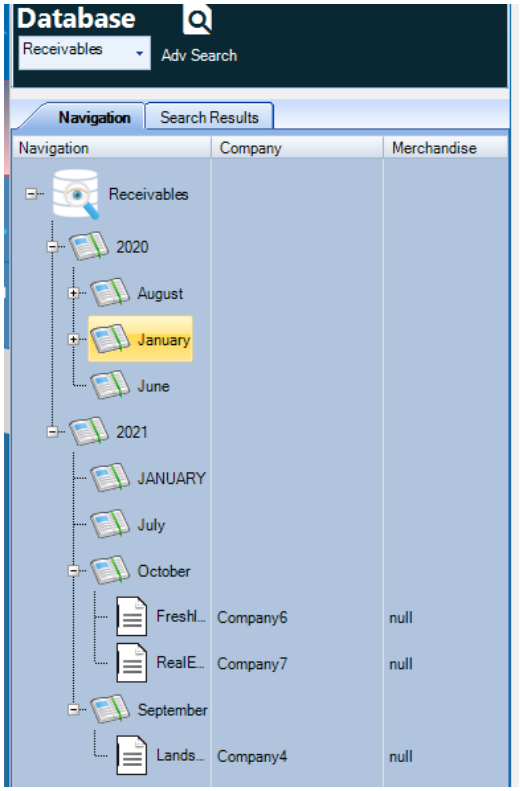
The search interface will present the available databases in a drop down field. Select any of your databases to begin searching
database found by default.



Once loaded, the Navigation panel will display multiple levels of trees based on any fields that are set as drop down items. An
will display as a tree level for easy visual navigation. Any fields that are text based are displayed in Index Fields as columns.
The document viewer will display any document that is selected and includes tools to pan, zoom and print the document.

When you click on a navigation item (IE A.R. Perry Glass), it will load the next level of drop down fields for that specific company type or display the documents for you. In this database setup, we only have the one field set for navigation (Company Name) so documents will display when we click a company name.





If we had configured two fields (Year and Month) to be dropdowns, we would have two le the Receivables database.

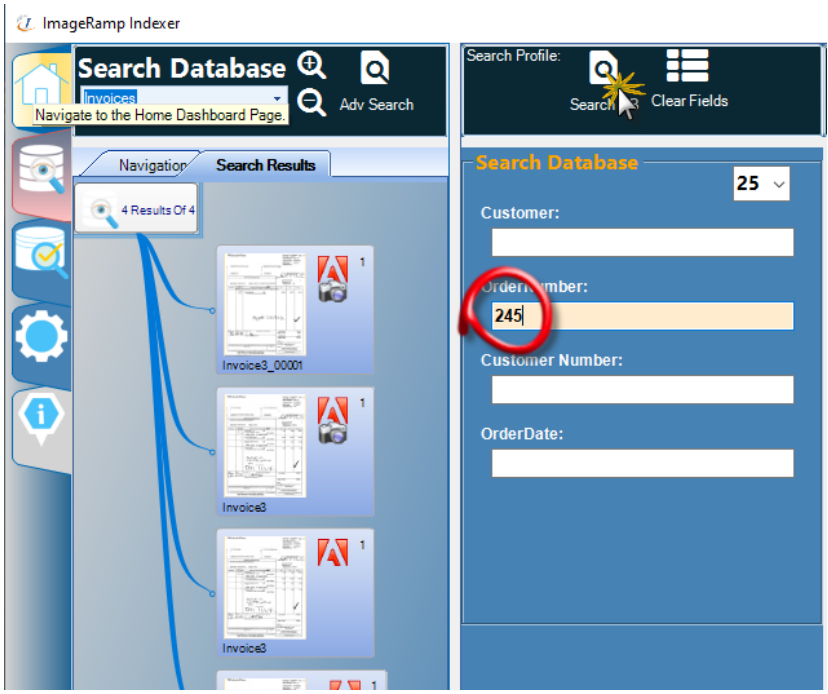
If you click on a year, all unique months found for that selected year are presente month, any documents matching the year and month are displayed in the tree.

Note: the order of the tree is based on the DisplayOrder in the field configuration

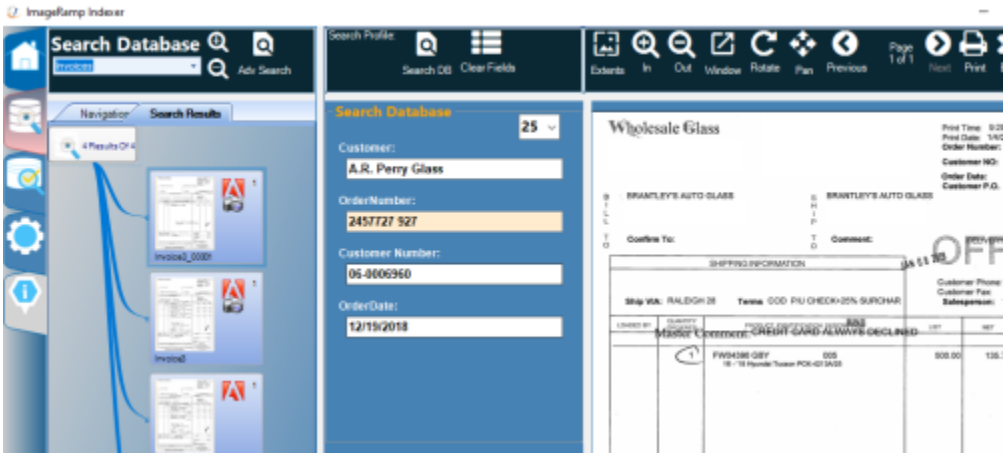
Using Advanced Search

If you want to search based on any keyword entry, use the Advanced Search. Selecting this button will invoke a panel that contains e configured fields. Simply enter a value in any field and it will look for all records containing that search string in that particular field. If you enter multiple entries, both conditions will be required for it to display resulting files.

In the example below, we are searching for any orders with "245" in that field. Simply enter the text and click the Search DB button and the resulting files will be displayed.



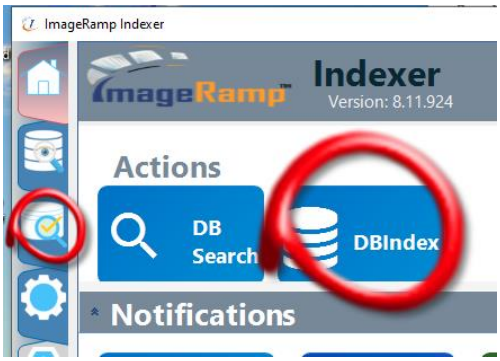
If you click on a document, the resulting file



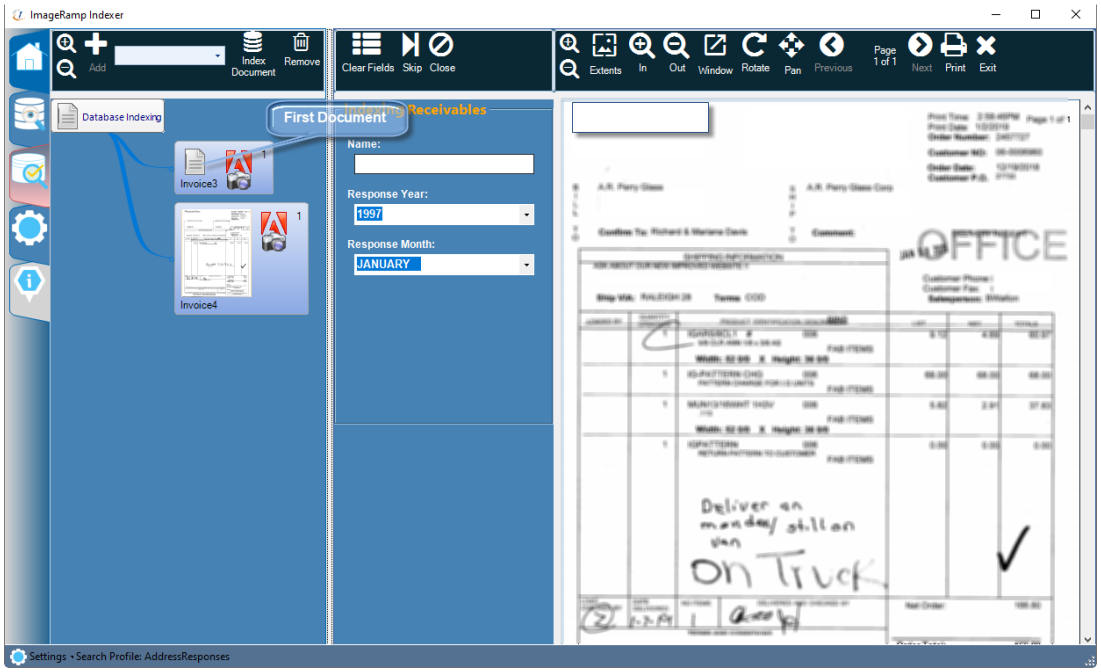
Database Indexing

Indexing Documents

If you want to add documents into the database, you would use the Indexing function. This is invoked by selecting the left pa
DBIndex tile.

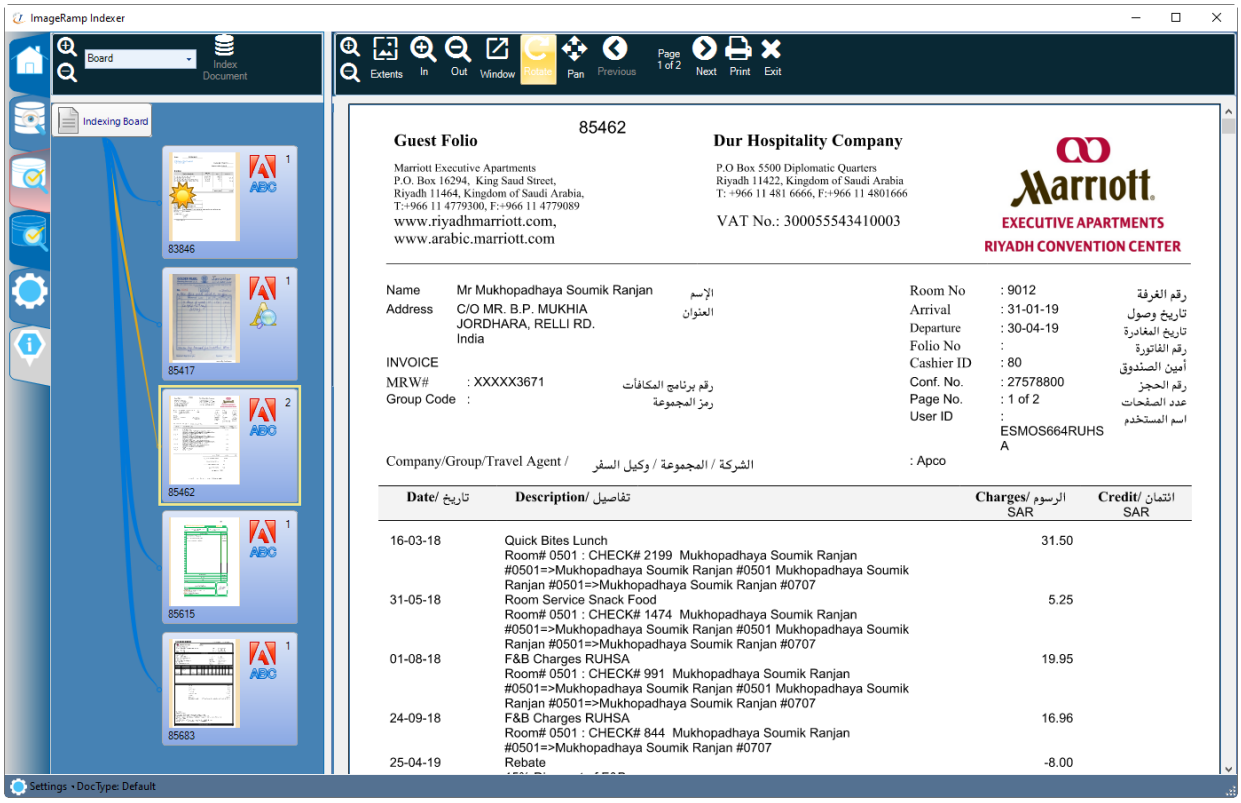


The user interface presented contains a tree listing of all documents contained in the datab
document is represented in an indexing node. The center panel is where the user would en



This is the main interface for indexing. The left panel contains the documents found in the sour doctype selected from the drop down. The center panel is the indexing data displaying any defa right panel is the previewer.

IF a user selects a document in the left tree, it allows them to preview their documents. The inde disabled and they can preview the pages associated with each new document to validate it is rea To index, they select the drop down again and the index button becomes enabled again.

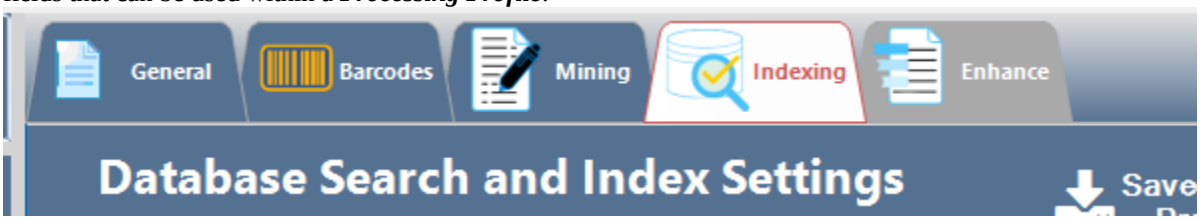


The indexing setup is fairly consistent with what it was before. I only added locations for for indexed documents. Index profiles are different than doctypes.

ImageRamp Settings – Indexing

Document Indexing Settings Overview

In this section, we address how to set up indexing in an ImageRamp ***Index Field Profile*** through the ***Indexing*** screen of the ImageRamp ***ImageRamp Profile Settings*** system. This Profile is not to be confused with creating a Processing Profile using the ***ImageRamp Profile Settings Menu*** of the ImageRamp interface. For more on the actual indexing of files and folders, see ***Initiating Index Processing***. In essence an ***Indexing Field Profile*** is a reusable indexing profile that maps fields that can be used within a ***Processing Profile***.



Indexing adds searchable information about your scanned files to a database or repository (third parties and Electronic Document Management (EDM) Systems, including DocuFi Online). Combining ImageRamp with a database provides the user with an effective means of scanning, organizing, distributing and hosting documents electronically.

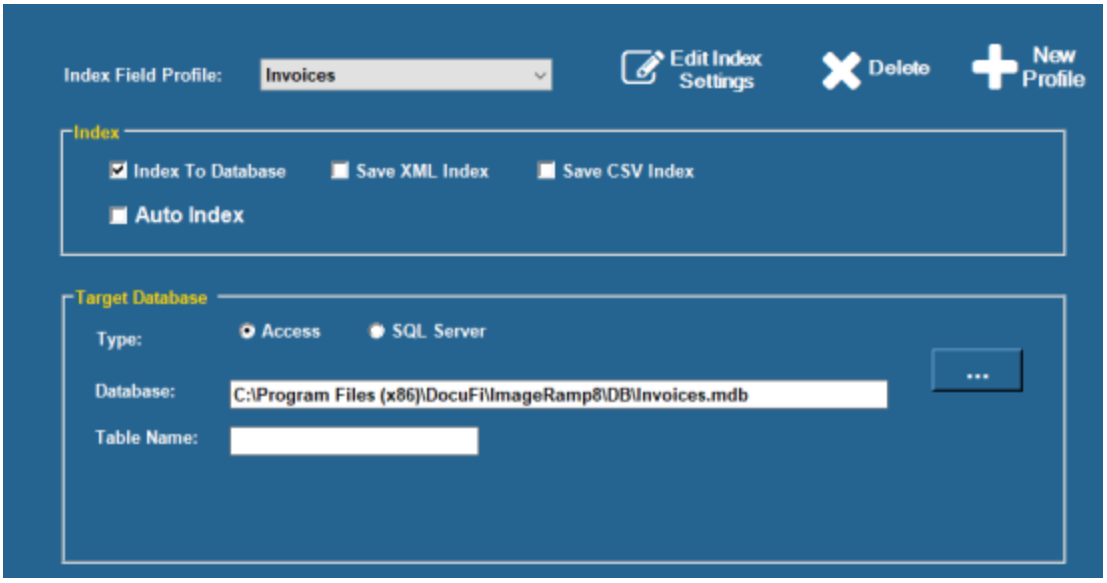
With indexing, apply categories of information to your scanned files. You can use common file properties, such as file name and date, and add customized fields captured from barcodes or data mining such as “invoice number”, "project name", "customer", "purchase order number", or "vendor". You can then search and sort these index fields when they have been incorporated in a database.

With ImageRamp’s simple interface and keyword codes it’s easy to add indexing information for your scanned files directly into a document management database or save the indexing information into eXtensible Markup Language (.xml) format or comma separated values (.csv) format.

When indexing to a database, setting up the index is a simple two-step process. We recommend using one of the pre-existing setups and adding and editing fields within them.

If you want to create your own ***Index Field Profile***, click on New Profile. This creates a template or map of the fields to index. Next, the selected ImageRamp Profile must be set to associate the proper ***Index Field Profile*** with the proper database or XML or CSV file. Depending on the target of the index (SQL Server, Access, XML etc.) different information is required.

Illustration: The Indexing options for ImageRamp Profiles.

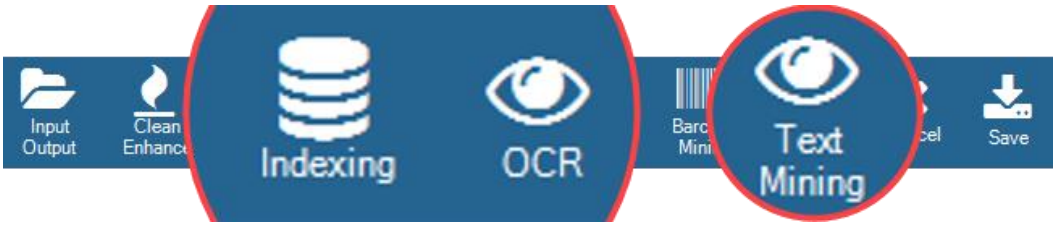


NOTE: Indexing via text mining and XML is only available with the Pro Edition.

NOTE: The ***Index Field Profiles*** are not to be confused Processing Profiles created using the ***ImageRamp Profile Settings Menu*** of the ImageRamp interface. ***Index Field Profiles*** simply contain settings for only indexing information.

NOTE: The ***Index*** options are not mutually exclusive; you may select multiple items such as both ***Index to Database*** and ***Save CSV Index***.

A Side Note about the Indexing, OCR and Text Mining Screens



To better clarify where related settings are located on ImageRamp screens, please refer to this chart.

Indexing	With the ***Indexing*** screen, designate what type of index you want to create such as XML, CSV, or index to a specified database. This screen is also used to set up ***Index Field Profiles*** which are reusable templates to create the index fields and their attributes. Optionally, it also can establish zones to apply OCR for implementing text mining to populate the index fields.
OCR	The ***OCR*** screen tells ImageRamp what OCR engine use for searchable PDF files and text mining; what advanced enhancement functions to employ to improve OCR accuracy, how fast to implement the OCR; and where it should look for data mining index zones; (if any have been created) to speed up the processing.
Texting Mining	The ***Text Mining*** screen allows you to use regex coding against a document's OCR results to create indexes; name files, folders and subfolders; and create document splits from mined text. The ***Text Mining*** screen selections determine how the OCR text data is extracted and used once the OCR engine has translated the image into text.

Managing Index Field Profiles

The ***Indexing*** settings from ImageRamp Profile allow you to create, configure and manage ImageRamp ***Index Field Profiles*** which are used when indexing information into a database, an XML file or a CSV file. Profiles are templates or maps of the fields to index. In these templates, users specify the field characteristics such field name, field type, default values and more.

Illustration: The Index Field Profile Screen with the pulldown menu of existing Index Field Profiles. The Invoices Indexing Profile is selected.

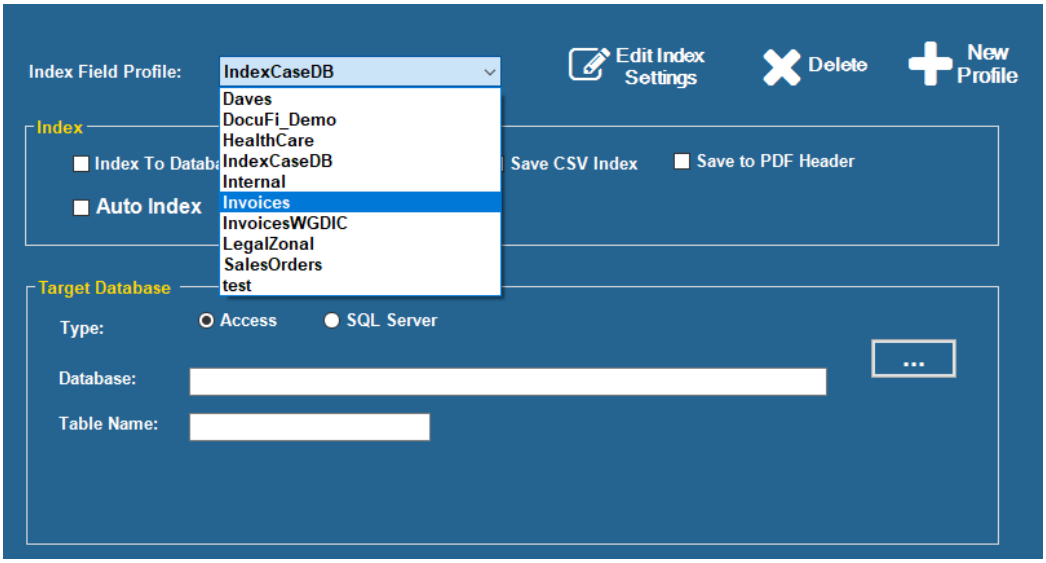


Illustration: The Index Fields of the Invoice Index Field Profile are displayed for editing when the Edit Index Settings icon is selected.

NOTE: The ***Index Field Profiles*** are not to be confused Processing Profiles created using the ***ImageRamp Profile Settings Menu*** of the ImageRamp interface. Index Field Profiles simply contain settings for only indexing information.

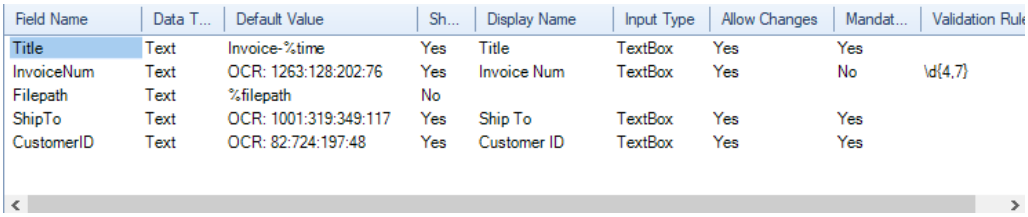
Creating a New ImageRamp Index Field Profile

To create a new ***Index Field Profile***, simply click on the **New Profile** button of the ***Index Fields Profile*** screen and enter a name in the pop up screen and click **[OK]**. You are returned to the ***New Field Setting*** screen where you can now enter the appropriate index fields for the new profile.

Editing ImageRamp Index Field Profiles

To edit an existing ImageRamp ***Index Field Profile***, load the desired ImageRamp profile, click on the **Indexing** icon, use the pulldown menu to select the desired ***Index Field Profile*** and click on the **Edit Index Settings** button. This displays the ***Edit Index Field Settings*** screen.

Illustration: Edit Index Field Settings screen



Field Name	Data T...	Default Value	Sh...	Display Name	Input Type	Allow Changes	Mandat...	Validation Rule
Title	Text	Invoice-%time	Yes	Title	TextBox	Yes	Yes	
InvoiceNum	Text	OCR: 1263:128:202:76	Yes	Invoice Num	TextBox	Yes	No	\d{4,7}
Filepath	Text	%filepath	No					
ShipTo	Text	OCR: 1001:319:349:117	Yes	Ship To	TextBox	Yes	Yes	
CustomerID	Text	OCR: 82:724:197:48	Yes	Customer ID	TextBox	Yes	Yes	

This screen is an entry screen similar to a table screen in MS-Access or a spreadsheet in MS-Excel. The entries into the table are controlled with the icons in the upper right corner and the up and down arrows.

With ImageRamp, you can define the following:

- Field Name – the source column name within your database
- Data Type – the type of index field to capture data (Text, Numeric, Date, etc.)
- Default Value – the initial value or keyword to use for data entry. Several keywords are available including barcode values, zonal OCR, file characteristics, date, time and station information.
- Show UI – toggles if the field is to be made visible to the user. Some fields may be necessary such as file date/time/path/extension that can be hidden but automatically captured.
- Display Name – the string we display to the user.
- Input Type – indicates if a dropdown box or text box is to be used for data entry.
- Allow Changes – indicates if users are allowed to modify values in a dropdown or text box.
- Mandatory – indicates if null (empty) values can be specified or if data is required before indexing is accepted.
- Validation Rules – employs technology to apply rules against the extracted information to confirm the validity of the content. For instance if an inventory item should contain three alpha characters followed by five numbers, all documents with item numbers that are not identified with that scheme during the capture process will be ignored. The user can then manually enter the proper data.

Managing Index Fields within an Index Field Profile

Adding and Deleting Index Fields

To add a new index field, select the profile from the pulldown menu, click the **Edit Index Settings** icon and then click the **Add** icon to bring up the ***Add a New Field*** screen.

In the ***Add a New Field*** screen, type the desired information or use the pulldown menus to select the appropriate information. Information includes:

Field Name	This is the field name in the database you are integrating into.
Data Type	Choices include: Text, Numeric, Date, Unique GUID Short for Globally Unique Identifier GUID is a unique 128-bit number that is produced by the Windows OS or by some Windows applications that can be used across all computers and networks wherever a unique identifier is required. Some database administrators

	even will use GUIDs as primary key values in databases.
Default Value	The Default Value is what is displayed if no input is provided. Keyword codes used in ImageRamp file naming may be used here as well as these codes: %FileName %bar1…….%bar*n* %FilePath OCR:xxx,yyy,width,height %FileSize %time %FileType %date
Show UI	Selecting Yes in Show UI (Show in User Interface) causes the item to be displayed in the ImageRamp Index Data screen at the glass. Selecting No hides the item from the user and is used for items like date, time, filepath, etc.
Display Name	This is the name which is displayed for the field in the ImageRamp Index Data Entry Screen
Input Box Type	This defines the type of box which is displayed for the user in the Index Data screen if the administrator has selected to Show UI. The options include: Text Box, Dropdown Box. The dropdown box is automatically populated with all the unique entries in the database when the user is presented with the indexing screen at the ImageRamp interface.
Allow Changes	Select Yes lets the user modify the data. Entering No means the user cannot change the field data.
Mandatory	Selecting No means ImageRamp will proceed and accept a blank field. Selecting Yes means the user must enter information before the file can proceed.
Allowable Characters	This field allows you to enter rules for content validation using Regex See Field Validation using Regular Expression Scripts (REGEX).

Save your changes by clicking the **Save** icon.

To delete a field from an existing index profile, select the profile from the pulldown list and click the **Edit Index Settings** icon. On the resulting ***Edit Index Field Settings*** screen, select the field by clicking its name and clicking the **Delete** icon. Save your changes by clicking the **Save** icon.

Modifying Existing Index Fields

To edit a field in an existing index profile, select the profile from the pulldown list and click the **Edit Index Settings** icon. On the resulting ***Edit Index Field Settings*** screen, select the field by clicking on its name then the **Edit** icon. This presents the ***Edit Field*** screen where you can enter the desired changes. Save your changes by clicking the **Save** icon.

Reordering Index Fields

Use the up and down arrows on the side of the ***Edit Index Field Settings*** screen to reorder the index fields as they appear to the user. Simply highlight the desired index field in the table and use the up or down arrow to change the order which is the order it will appear to in the output such as a CSV file or XML file.

Click on the **Save** button to save the index profile changes or the **Cancel** button to cancel the index profile changes.

Field Validation using Regular Expressions - Allowable Characters

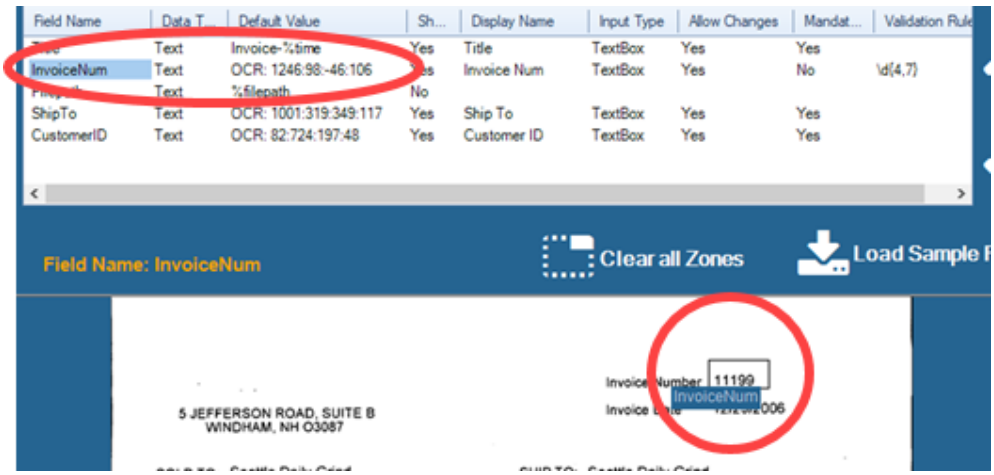
ImageRamp can capture field data and take advantage of regular expression scripts to validate the data is correct for indexing. You can define any character pattern that is to be allowed in the input string but this requires a knowledge of Regular Expressions (regex). If the entry does not match the script, the operator is given a warning message.

Please refer to the text mining panel and the regex library.

Indexing with OCR and Zones

Pro Edition users can set OCR zones on documents by first loading a sample file. Click on **Load Sample File** and browse to your sample document that represents your document set. Select one of your fields listed in the ***Index Fields*** list. Move your mouse to hover over the document. Your mouse will now display the field name and a rectangular select tool is presented. Select the desired area on the sample document. Once the area is selected, it is stored in the fields default description as an OCR zone.

Illustration: Setting an OCR Zone in ImageRamp



You can now reposition or enlarge the area using standard move and resize mouse controls. Just select the item and move or resize.

NOTE: OCR is only available with the ImageRamp Batch Pro Edition

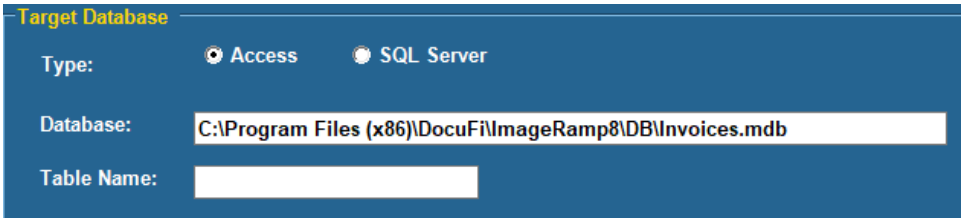
Index Settings for Access and Excel

Once the ***Index Field Profile*** has been selected, simply select to index to either a SQL Server or Access database, index to XML or a CSV file. For database indexing, authentication against the

SQL server instance will be required including then name of the database and table name to receive the index information. The table default is “documents” and may be changed by advanced users for their own specific document management environment.

Several sample databases are provided in the installation folder located in the DB sub folder. ?

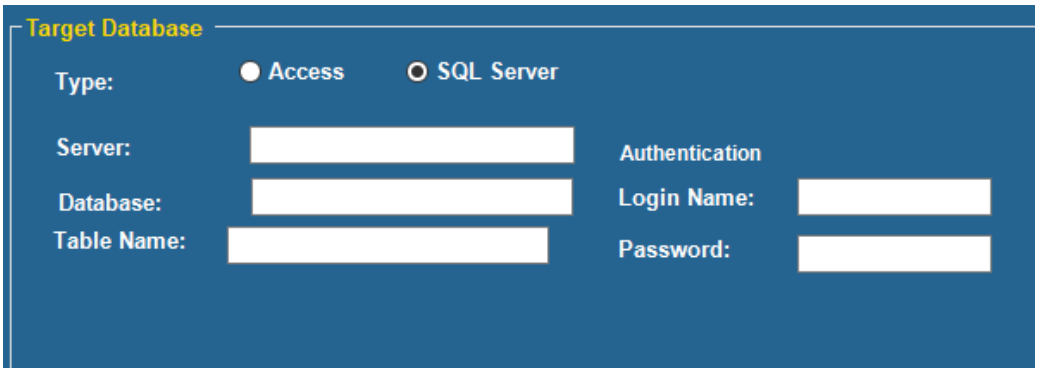
Illustration: Indexing to Access



Index Settings for SQL Server

Once the ***Index Field Profile*** has been selected, simply select to index to SQL Server; name the server, database and table; and enter the authentication information.

Illustration: Indexing to a SQL Server Database



Indexing to XML, CSV or PDF Header and Auto Index

XML was created to structure, store, and transport information and is a common integration format used with many databases. CSV files are often used as a simple way to transfer a large volume of spreadsheet or database information between programs without worrying about special file types. Both methods can be used to index into most commercial document management solutions. The CSV and XML files are placed in the same folder as the image.

To save the index information in either of these formats, simply check the ***Save XML Index*** or ***Save CSV Index*** boxes in the Index area of the ***Index*** Screen.

Illustration: Indexing to a Database via XML or CSV

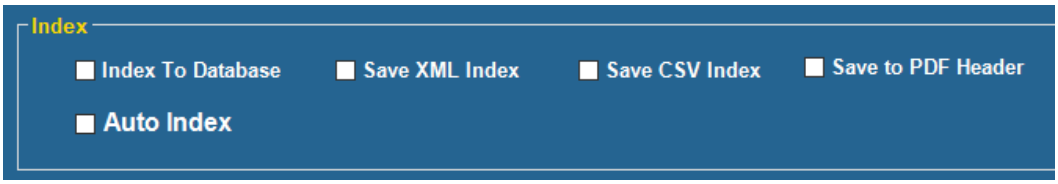


Illustration: Results of saving the index as XML from ImageRamp sample

<?xml version="1.0" ?>

– <SPIndex>

<index1 name="**Title**">**888-2008-07-17-085002669.tif**</index1>

<index2 name="**Customer**">**DOE,JOHN**</index2>

<index3 name="**Document_Type**">**888**</index3>

<index4 name="**Customer_PO**">**DB43868Z**</index4>

</SPIndex>

Checking ***Save to PDF Header*** adds the selected index information to the PDF file header and allows the user to implement advanced document manage functions using the information.

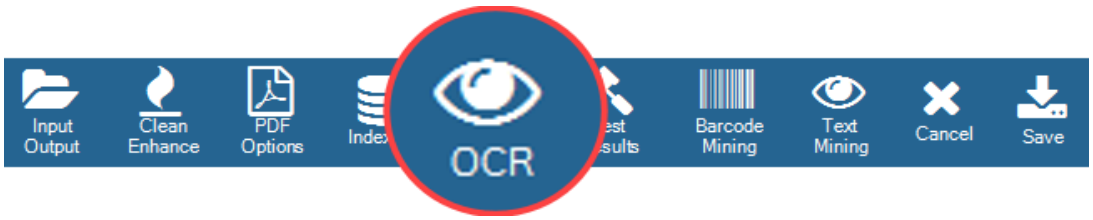
NOTE: The ***Index*** options are not mutually exclusive; you may select multiple items such as both ***Index to Database*** and ***Save CSV Index***.

NOTE: See Initiating Index Processing to understand how automatic and manual indexing processing is done.

Auto Indexing vs. Interactive Indexing

When index to XML, CSV, a database or PDF Header is selected, ImageRamp allows you to step through each document manually and verify or edit the index information in a popup window titled ***Folder Indexing*** when the processing is initiated. If you desire to automatically index all the documents in a folder, select the ***Auto Index*** checkbox. When ***Auto Index*** is selected, the ***Folder Indexing*** window for manual viewing/editing is not shown unless an error occurs. See Interactive Index Processing for more information.

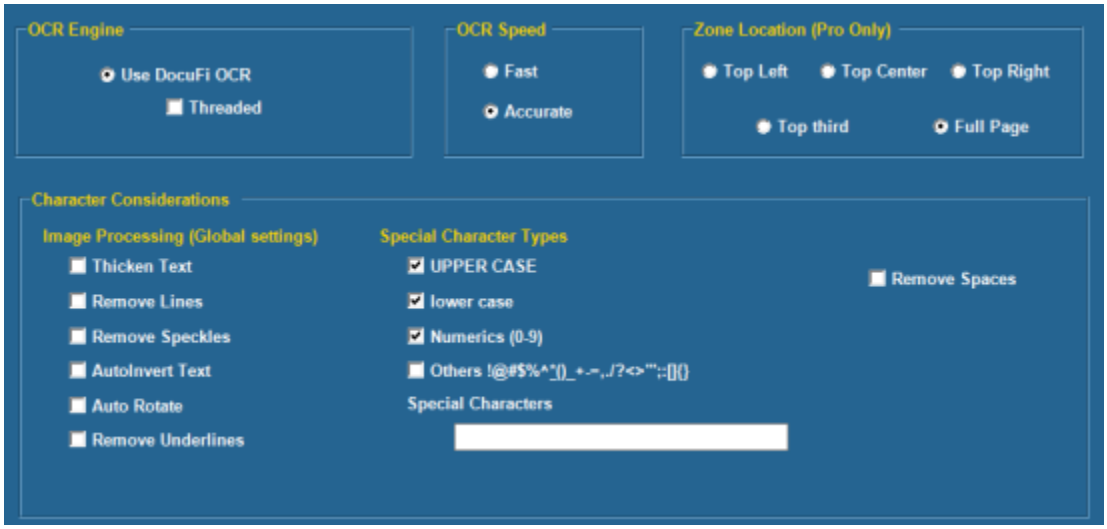
ImageRamp Settings – OCR Options



Setting OCR Options

If you have elected to output your files or scans as ***PDF-OCR*** (searchbe PDF) on the ***Input Output*** screen or want to use text mining for indexing, file/folder naming and document splitting, ImageRamp Batch allows users to determine how the OCR engine is applied. The OCR options apply for indexing with zonal OCR and drag-and-drop OCR in manual processing. OCR is available in the Pro Edition. See A Side Note about the Indexing, OCR and Text Mining Screens for further clarifiction on the impact of the OCR settings.

Illustration: The OCR options interface.



Zone Location For Indexing

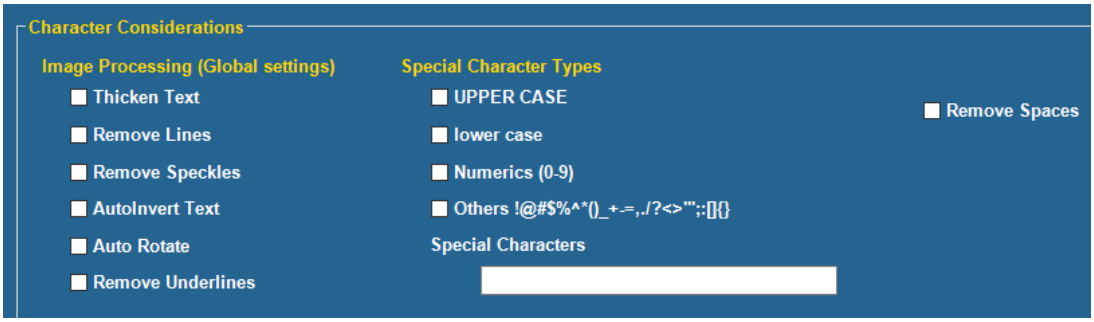
Scanning speed can be greatly effective by narrowing the OCR location. Select the area to OCR to collect the desired index information by simply selecting the area in ***Zone Location***.

Characters for Zonal or Drag-and-Drop Drop OCR

The OCR settings are tailored to improve OCR accuracies and take advantage of text data to help determine AutoRotation and other functions. These options are applied to the OCR processing in addition to any enhancement processing identified in the ***Cleanup Enhance*** screen.

Character enhancements are as shown in the illustration

Illustration: Character enhancement options.



The special characters allow you to specify exactly which character types to look for in a document. If you only want characters and numeric for instance and ignore special nomenclatures (#@%$..) then this can be accomplished by selecting the appropriate entries.

Remove Spaces will remove all space characters from the resulting data. This is used mostly with zonal or regional data mining purposes.

You can specify specific characters by entering those characters in the ***Special Characters*** text box.

TIP: Pretest your sample scans using ImageRamp's preview function as described in ImageRamp Settings – Test Results to determine which OCR engine and speed work best for your documents.

Interactive Index Processing

Initiating Index Processing

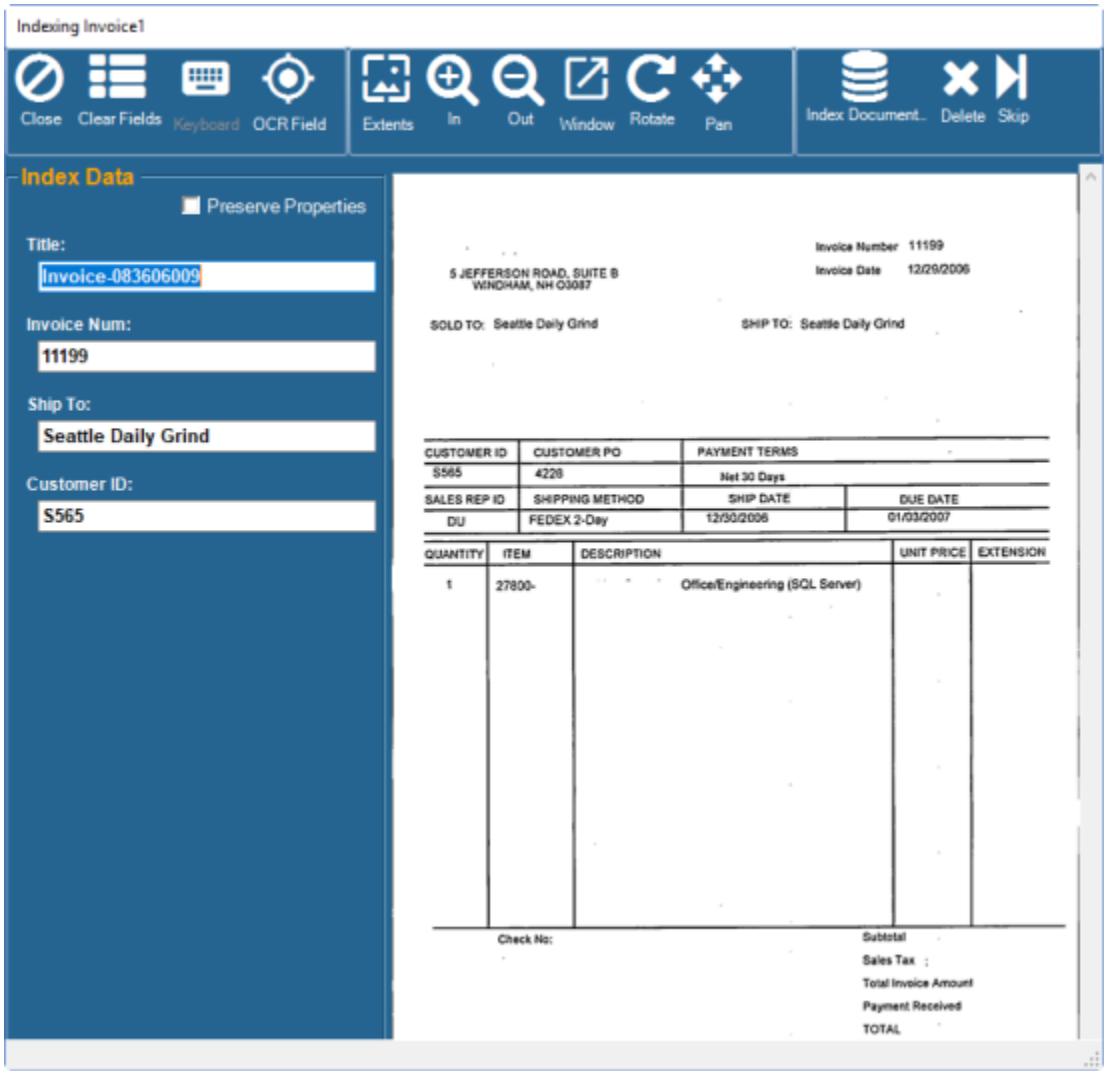
Once the **Indexing** settings have been selected in an ImageRamp profile, ImageRamp can easily create index information about files to store into a database, XML or CSV file. See ImageRamp Settings – Indexing to learn how index fields are created.

To manually inspect and edit what information was automatically captured based on a profile's settings, users can start an index-enabled profile and start processing files. Each unique file will be loaded and data mined and presented into the designated field within the ***Index Data*** interface.

The ***Index Data*** is presented as set up in the ***Index Field Profile***. In the example illustrated below, four fields (Title, Invoice Num, Ship To, and Customer ID) were populated with data from OCR Zones set up in the indexing profile. Note: barcodes and other data can also be incorporated into the indexing process. See Indexing with Text Mining and Barcode Mining.

This process offers a unique ability to QA the data in an efficient manner. The mined data is presented and allows operators to validate or replace the data. Tools are available to view the document closely such as zooming, pan, etc.

Illustration: The indexing results in a loaded file after indexing has been setup in a profile.



Once the index information has been entered, the file can be processed and stored into the database, xml or csv packet by clicking **[INDEX DOCUMENT]** or skipped by clicking **[SKIP]**.

NOTE: Files processed by the indexing process will be removed from the source folder. It can now be processed or skipped and the next file in the folder is then presented.

Interactive OCR Indexing (Drag-and-Drop Indexing)

Often users want to replace the text or manually enter new text into fields not assigned to barcodes or OCR zones.

To OCR information into an index field:

1. Start processing an index-enabled profile,
2. Place the cursor in the desired index field in the left index field pane and select it,
3. Click on **[OCR FIELD]**, and relocate to the document area. The mouse will have a field name shadow it as you move about the document
4. Draw a rectangle around the desired text in the right document viewing pane by holding the left mouse down and dragging the box around the desired text.

ImageRamp automatically OCRs the boxed text and places it in the index field box. Users can edit the text within the index field box if needed.

Illustration: The OCR'ed text in the index field.

NOTE: Once documents are indexed, they are removed from the source folder when process folder is utilized.