



SimpleIndex for Invoices Set-up & Instructions

Welcome to the getting started guide for SimpleIndex for Invoices. Capturing data from documents (scanned or electronic) and extracting the key information. With a brief set-up process SimpleIndex can automatically read the Vendor Name, Invoice Number, Date and Total Amount, along with some other Header and Footer information from invoices that are in different formats from different sources.

The first thing that should be done is to download the SimpleIndex demo from the [Download Page](#)

Once the software has been downloaded and installed go through the Global Settings Wizard after installation and make sure that “Copy sample jobs to folder” is checked. This will copy all of the Sample Jobs included with the SimpleIndex download to the Default Job Configuration Folder, which is on “C:\Images” by default, but can be changed to any folder required that the computer SimpleIndex is installed on has access to.

The Sample Jobs show several typical processing workflows and are the best way for new users to get

Default Job Configuration Folder

c:\Images

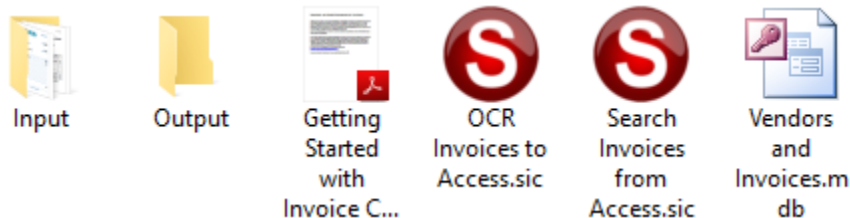
Browse...

☒ Copy sample jobs to folder

Once the Global Settings Wizard is complete all of the settings files for this process can be found in this folder:

“C:\Images\SimpleIndex for Invoices\SimpleIndex for Invoices Standalone”

This folder contains the following files:



- Input – This is the folder that this process will import images from. Any images in this folder will be processed with the Job Configuration. This folder contains some sample invoice images that have already been configured to recognize with the demo.
- Output – This is the folder where the invoice image files will be saved once they have been fully processed.
- Getting Started with Invoice Capture – This is the PDF file with the instructions for how to configure and run the Invoice Processing Job Configuration.
- OCR Invoices to Access – This is the SimpleIndex Job Configuration file that would process the images, recognize the key information and save the images to the Output folder and the indexes to the Database.
- Search Invoices from Access – This is the SimpleIndex Job Configuration file that would allow searching for the invoices that have been processed with the OCR Invoices to Access Job Configuration based on the Vendor, Invoice Number, Date and Total Amount.
- Vendors and Invoices – This the Access Database file that contains the table that is used to read the information from the invoices and also the table that stores the exported index information.

To run the sample invoice images included with the demo download follow the instructions below.

1. Open SimpleIndex
2. Click the Open Job Button



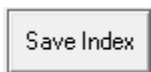
3. Open the “OCR Invoices to Access.sic” from the “C:\Images\SimpleIndex for Invoices\SimpleIndex for Invoices Standalone” folder.



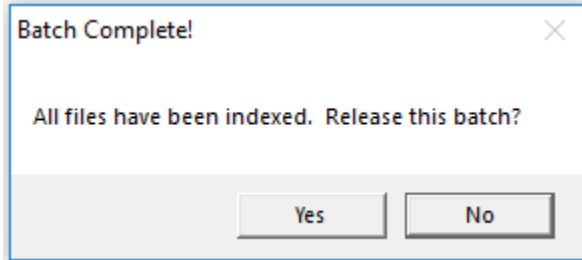
4. Click the Run Job button.



5. Once the Run Job button is clicked the Job is started, which imports the images from the Input folder, does OCR on the page, extracts the key information and then displays the first image in the batch for review.
6. In the review process each image is looked at and the read index fields are confirmed correct. If not they each incorrect value is changed. After reviewing each page, the Save Index button is checked to move onto the next invoice.



7. Once Save Index has been pressed on each page in the batch of invoices when you click the Save Index button on the last page you will get the “Batch Complete!” window. When you click Yes, it will Export all of the processed images to the Output folder and the data to the database.

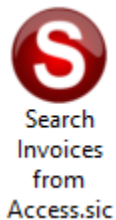


To search for the invoices scanned with the “OCR Invoices to Access.sic” Job Configuration file follow the instructions below.

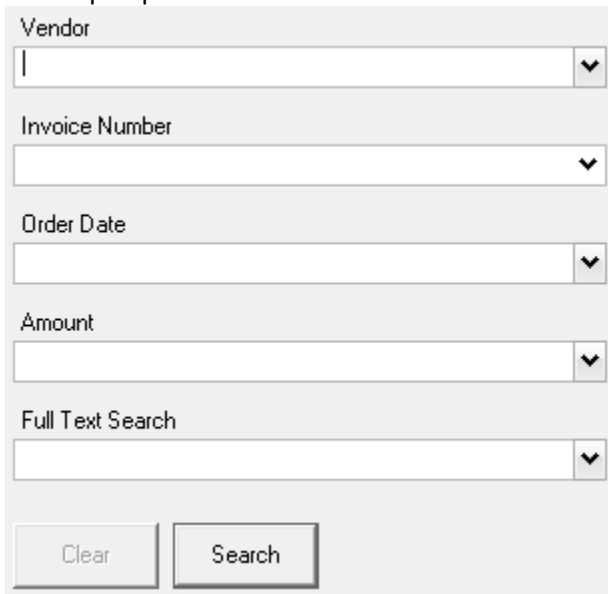
1. Open SimpleIndex
2. Click the Open Job Button



3. Open the “Search Invoices from Access.sic” from the “C:\Images\SimpleIndex for Invoices\SimpleIndex for Invoices Standalone” folder.



4. Once open pick one or more values in the search fields and click the Search button.



The screenshot shows a search interface with five dropdown menus stacked vertically. The labels for the dropdowns are 'Vendor', 'Invoice Number', 'Order Date', 'Amount', and 'Full Text Search'. Each dropdown has a small downward arrow icon on the right side. Below the dropdowns are two buttons: 'Clear' and 'Search'.

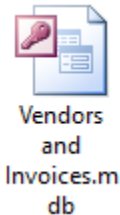
5. Once the Search button is clicked the invoices that were searched for will be displayed.

The SimpleIndex software can also be configured to recognize invoices that aren't a part of the samples included with the demo. The instructions for setting this up are below.

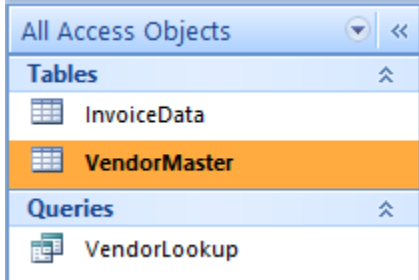
1. Put new invoices that need to be processed into the Input folder and move any sample image files that don't need to be processed out to another folder.



2. Open the "Vendors and Invoices.mdb" Access Database file, which requires a licensed copy of Microsoft Access. If a licensed copy of Microsoft Access is not available then there is a free tool that can be used for this process, which can be downloaded [HERE](#). The instructions from this point forward will all be based on Microsoft Access.



- Once the "Vendors and Invoices.mdb" Access Database file is open, double click the "VendorMaster" table to open this table to add new invoice templates.



- Once the "VendorMaster" table is double clicked this table will appear:

ID	VendorName	InvoiceNumberTemplate	VendorNameDictionary
19	Name of Vendor	Name of Vendor alternate phrase on page 2nd alternate	alternate phrase on page 2nd alternate
53	Tech Type Interfaces	Invoice#####	Tech Type Interfaces 123-456-7890
54	Paper Scan, LLC	InvoiceNo###AA###	987-654-3210 http://www.paperscan.com ap@paperscan.com
55	One Source Inc.	###-####-#	One Source Inc. 15 Rhondo Drive 456-123-7890 info@1source.net
56	Document Giant	INVOICE\#####	Document Giant 789-456-1230 purchasing@docgiant.com
57	Dictated Records Storage	INVOICE\#A#####A	Dictated Records Storage www.d-rs.com
58	Fire Fly Corp.	INVOICENO#####	456-987-1230 97 Holmes Blvd. invoicing@firefly.com
*	(New)		

- Next type in the Company Name of the invoice that will be configured first in the "VendorName" column next to (New). This will be the name that is used for this Company when saving the invoices to the Output folder and to the Database.

58 Fire Fly Corp.	IN
(New)	

- Next keywords need to be added to the "VendorNameDictionary" column on the same row as the new Company Name. These will be key words or phrases that are unique to the particular Company Invoice template that needs to be read.
 - Enter the first keyword or phrase that is unique to that particular Company invoice template. Some good values for this field are the Company Name itself, Phone Number, Address and Website. These are good because they will always appear on this Company's invoices, but never on any other invoices.
 - If one or more additional keywords or phrases are needed then they needed to be separated by a "|" (shift + \), which is also called a pipe or vertical bar. This should be entered immediately after the value without spaces and the next value should be entered immediately after the "|" without spaces. EX. Value1|Value2
 - As many keywords or phrases can be entered in the "VendorNameDictionary" as required, as long as each one is separated with a "|".
 - As many words can be entered per section and all of the words have to be together in the text of the page to match. EX. If you put the value "these words are on the invoice" then the whole phrase has to appear to together in order in the page text to match. If it sees "words are on the invoice" or "these words on the invoice" then it won't match because all the words in the phrase aren't there in order.
 - It is possible to have multiple words that appear in the OCR text, but they don't appear together in order. To do this separate the keywords or phrases with "&&". EX. Value1&&Value2&&Value3|Value5 In this example if all three of Value1, Value2 and Value3 are anywhere on the page it will read as that record.
 - Is it possible to do a negative keyword match. With a negative keyword match if a certain keyword or phrase appears in the text it will skip that record and move onto the

next. To do this put a “^” before the value will be used to skip the record. EX.
Value1|^Value2 In this example if “Value2” is in the text it will skip this record.

7. The Invoice Number Template is used to find the Invoice Number from the image based on a certain text format/pattern/template. Once the process has recognized the Vendor Name on the page it then applies the Invoice Number Template to a search against the text on the page to find the Invoice Number.

- a. Once the “VendorName” and “VendorNameDictionary” columns for a new Company’s Invoice are created the template for the invoice number will be put in to the “InvoiceNumberTemplate” column on the row with the new Vendor that is being created.

VendorName	InvoiceNumberTemplate
9 Name of Vendor	Name of Vendor alternate phrase on page 2nd alternate
3 Tech Type Interfaces	Invoice#####
4 Paper Scan LLC	InvoiceNo####A###

- b. In the “InvoiceNumberTemplate” column you put one character in for each character in the pattern that the Invoice Number will be. The potential template characters are listed below.

= Number

A = Letter

X = Letter or Number

* = Any Character

EX. If the Invoice number is “123ABC” the “InvoiceNumberTemplate” would be “####AAA”

- c. Qualifying text can be put before or after the template for the value you want to read. This makes it possible to have the value looked for include the label on the page that comes before or after the value as well as the template. Once the value is read any qualifying text before or after the template characters are stripped out.

EX. If on the page the Invoice Number has the word Invoice before it and then the Invoice Number is “12345-00” then the template would be “Invoice#####-##”

- d. The following characters are automatically removed from the text that is read from the page: ; , Hard Return Space

- e. If the qualifying text has one of the template characters in it then you can put a “\” before the value so it reads it specifically. This is most useful when the qualifying text has the letter “A” or “#” in it, but can be used for any of the template characters that you want to read as the actual character.

EX. The Invoice Number on the page is “Account Invoice # DD1234-C” the “InvoiceNumberTemplate” value would be “\Account Invoice\#AA###-A”

8. Once the “VendorName”, “VendorNameDictionary” and “InvoiceNumberTemplate” have been filled in for a new company it is a good idea to test one to ten invoices from that company to make sure everything is reading properly. If not then adjustments can be made and then tested again.

The final index data is saved in the “Vendors and Invoices.mdb” database in the “InvoiceData” table.