



PSIGEN software inc.  
the **science** of document capture

QUICKstart

01

*Modified: 12.5.19*  
*Screenshots: 7.5.0.11*



PSI**capture**

# Smart Zones

# Quick Start Guide

## Getting Started

### *Welcome to PSICapture!*

**T**his Quick Start Guide is designed to help you through the process of creating a Capture Profile Configuration that uses Smart Zones to extract index data. This guide includes sample images that will be used to configure and test the Capture Profile. For additional information on features not covered in this guide, please refer to the PSIGEN wiki at [help.psigen.com](http://help.psigen.com). PSIGEN offers detailed hands-on training with the PSICapture software as well.

Make sure you have completed the **00 - Installation Quick Start Guide** and used it to get PSICapture installed, licensed and running on your computer before starting with this Guide. You will need to be running version 7.6.x or newer for this guide to be fully effective.



#### **Requirements:**

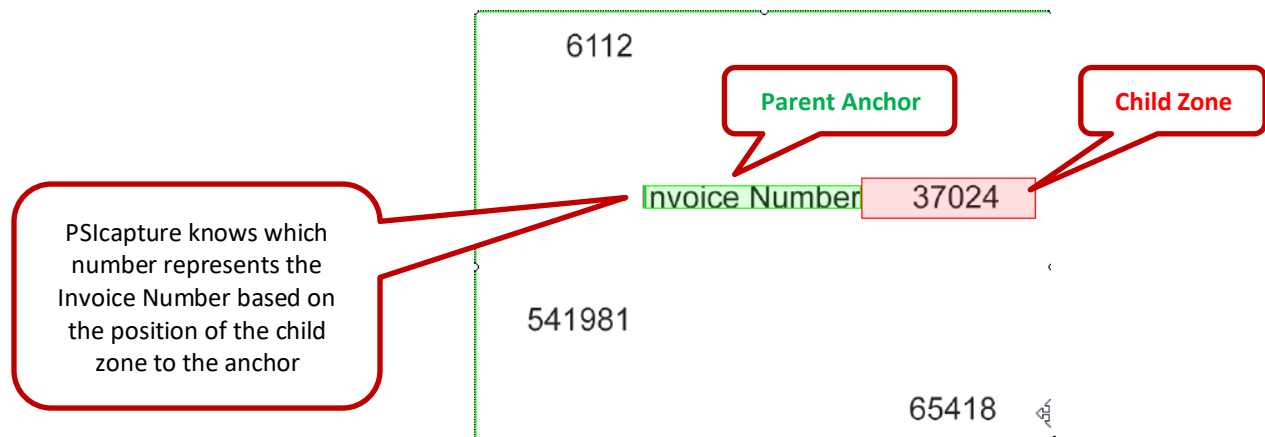
- Completed **00 - Installation Quick Start Guide**
- Ensure your computer meets the [min. requirements](#)
- Running version 7.6.x of PSICapture or newer
- PSICapture License with OCR and ADE modules

### What is a Smart Zone?

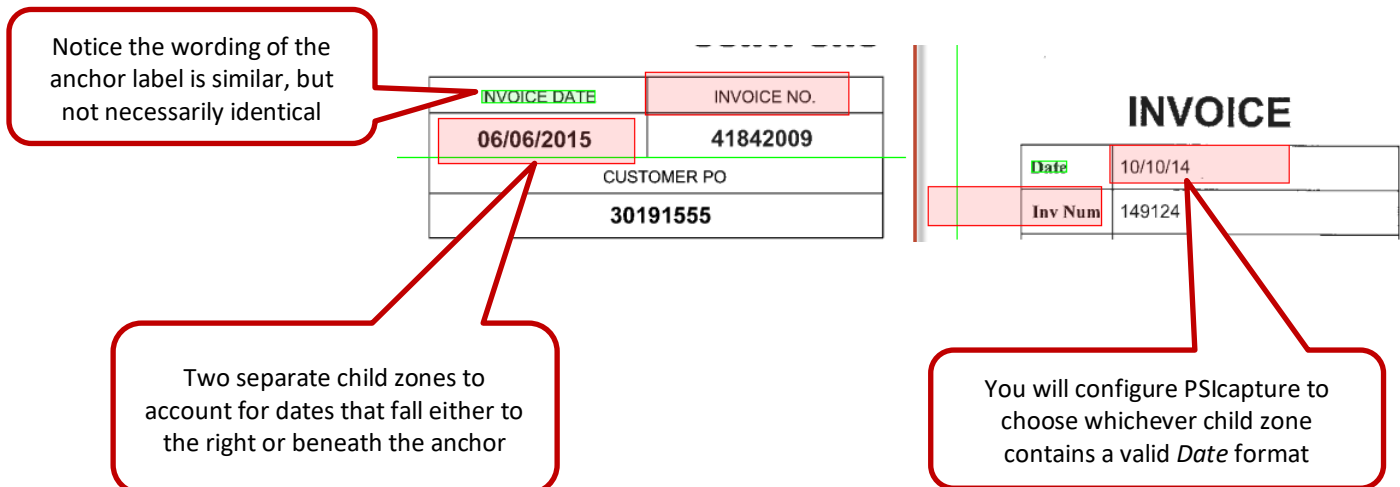
A Smart Zone is a type of zone from which PSICapture can extract index values. Smart Zones are especially powerful because, unlike Standard OCR Zones, Smart Zones can find index data even when it is not necessarily located in the same place on every document.

Configuring a Smart Zone first requires searching for an anchor. An anchor is a word, phrase, or other text value that appears in a position that is consistent in relationship to the index data you want to extract. Many times, the most convenient anchor will take the form of an adjacent label that identifies the data.

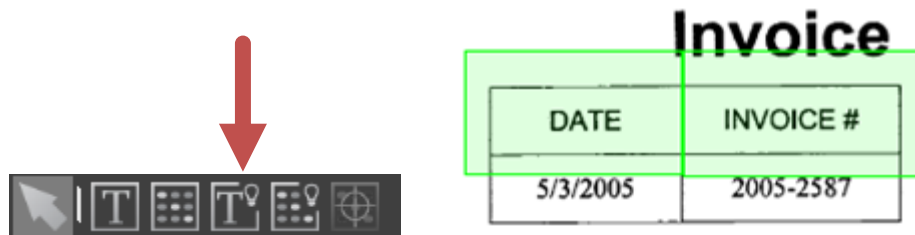
For example, on invoices, the Invoice Number often consists of just a series of digits. The Invoice Number will most likely be one of many numeric sequences found on the page. You can identify it as the Invoice Number because it is located in close proximity to the label words **Inv. No.** (or some permutation thereof). Smart Zones apply this concept to find and extract index data. Although the Invoice Number itself may move around on the page, you expect that it should be positioned consistently relative to the anchor. Each Smart Zone uses its own independent anchor.



Once the anchor is located, you can set up one or several child zones in which to look for the data you need to extract. If you know that the data you need will always be to the right of the anchor, you may only need one child zone. If the data could also be found beneath the anchor, then you may need 2 or more child zones. You can change the position and dimensions of each child zone to ensure they will contain the required data on all of the documents. Using regular expressions, you will specify the pattern of data to search for within the child zones. You will then configure PSIcapture to search each child zone in sequence, but only until it finds a string of text that matches the specified pattern.



Within the Zone Configuration window, the Light Bulb icon denotes the toolbar button that will allow you to draw a new Smart Zone. The Smart Zone should be drawn large enough so that it will contain the anchor on all documents. Smart Zones are easily recognized by their green color.

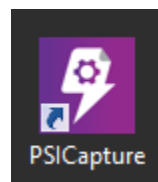


## How will the Capture Profile function?

You will build a single default zone profile that will extract index data to populate multiple fields using invoices from multiple vendors. The index data will be found in different positions on the page, depending on factors such as which vendor the document belongs to, and how many line items pertain to the document. Sometimes the data to be extracted will be found to the right of the label anchor, and other times it will be found beneath it. Despite these differences, PSICapture will be able to extract the correct values for every field on every document.

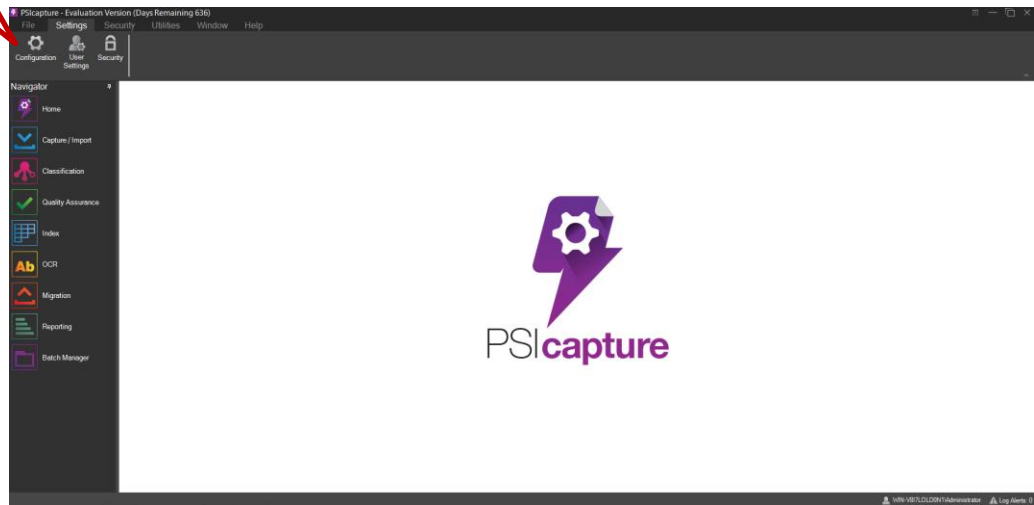
## Creating your Capture Profile

First, you'll need to get a basic Capture Profile Configuration started. Open PSICapture.

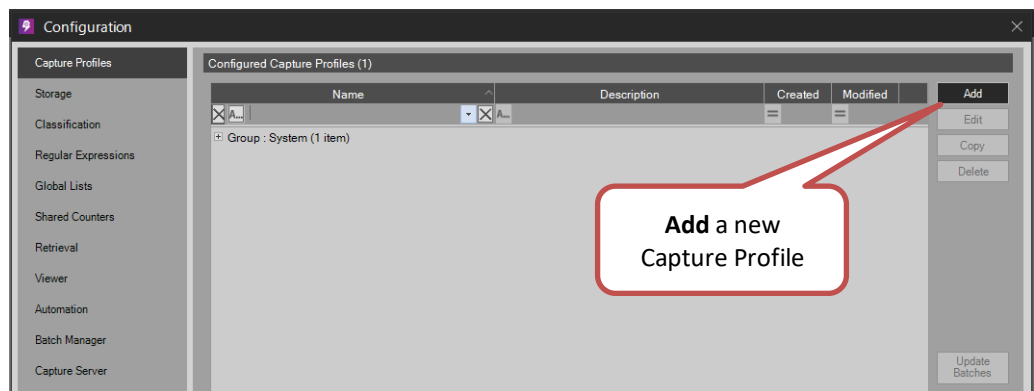


## SMART ZONES QUICK START GUIDE

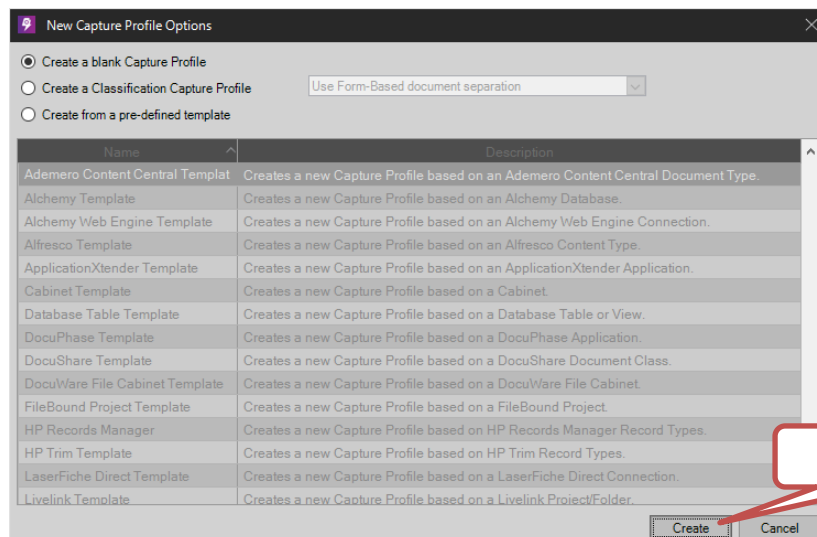
Select to open  
Configuration



After PSICapture opens you will start in the Configuration module to create the Capture Profile.



Once the Configuration module is open you can begin work on your first Capture Profile. Click the **Add** button on the right side of the window to proceed.



## Capture Profile Tab

The screenshot shows the 'Capture Profile Configuration' window with the 'General Setup' tab selected. The left sidebar lists various configuration categories. Five red callout boxes point to specific fields: 'Name' points to the 'Name' field, 'Group' points to the 'Group' dropdown, 'Description' points to the 'Description' text area, 'Batch Storage' points to the 'Batch Storage' dropdown under 'Storage options', and 'Batch Retention' points to the 'Batch retention type' dropdown under 'Batch retention'. The 'Batch retention' section also includes radio buttons for 'Immediately' and 'After (days):' with a numeric input set to 5.

Callouts in the image:

- Name
- Group
- Description
- Batch Storage
- Batch Retention

Give your new Capture Profile a name, group, and description. Confirm your Batch Storage Location and set your Batch Retention.

## Auto Import Tab



For the purpose of this Quick Start Guide you will not be making any changes to Auto Import or it's sub tabs. Many of the features within PSICapture Capture Profile Configurations can be left at their default settings. For additional information about features available on this tab please refer to the PSIGEN [Help](#).

## Document Records Tab

Again, leave all the default settings for Document Records, as they are not needed for this Capture Profile Configuration.

☆ Note:

Separation profiles are unique to each Capture Profile.

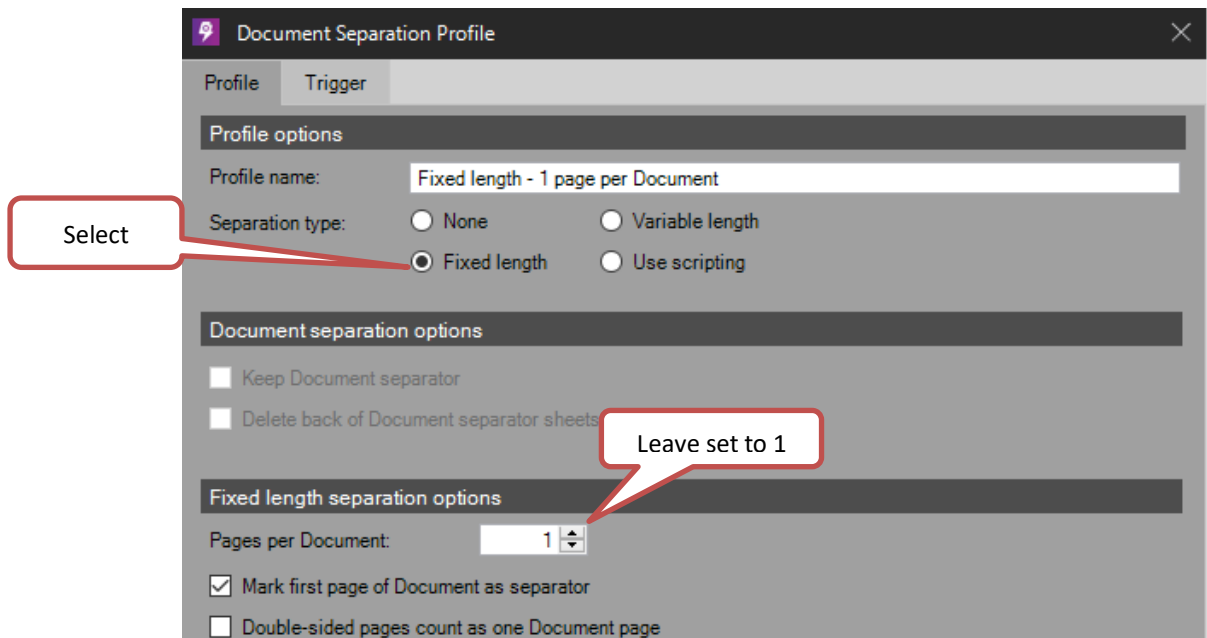
## Separation Tab

This is where you will configure a Document Separation profile for the Capture Profile. There are many options for document separation and this guide only covers one of them.



Please select **Add** to create a new Document separation profile.

## Document Separation Profile



Select Fixed Length separation and leave the default of 1 page. If you review the sample documents that you're using, you will see that each invoice consists of only one page. Click **Save**.

## Recognition Tab

No barcode or other recognition engines are being used in this guide, so you can skip this tab.

## Indexing Tab

Here you will configure the Document Index Fields.

Click the **Add** button on the middle of the window 4 times to create 4 blank index fields. Name those index fields as follows:

- Vendor Name
- Invoice Number
- Invoice Date
- Invoice Total

Click **New** 4 times  
(once for each field)

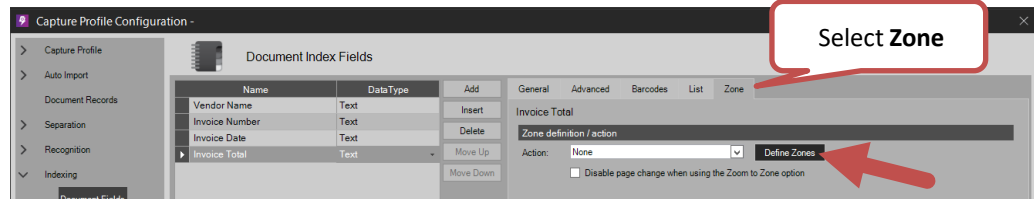
Notes

Apply Save Cancel

Now you've got the index fields defined. Next, you will create the Smart Zones from which to extract the data to populate those fields.



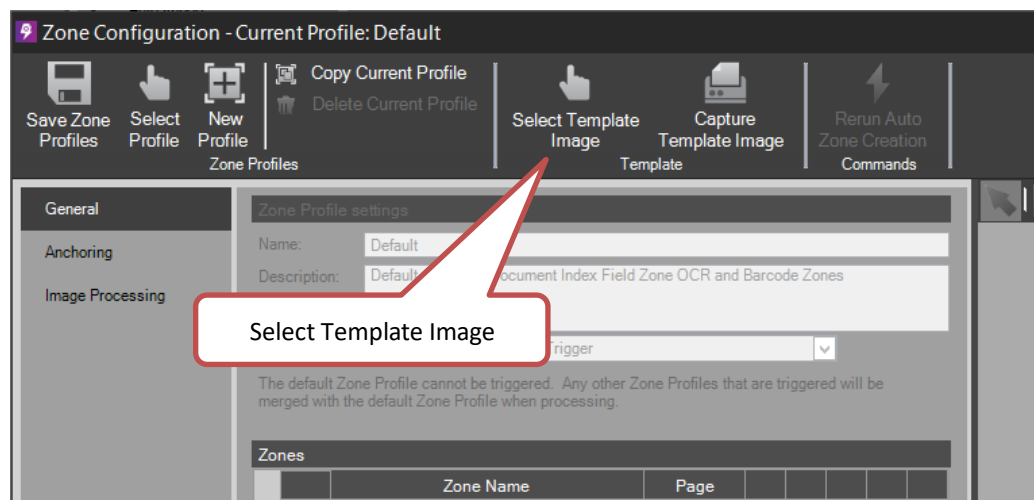
Click on the **Zone** tab and click the **Define Zones** button.



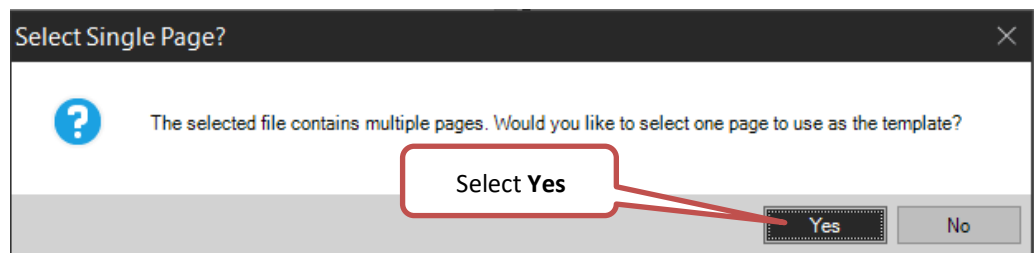
☆ Note:

It is a best practice to evaluate all sample images before you draw your zones, in order to verify that your placement will work for all documents.

You need to load in some samples for the image types you will be processing with this Capture Profile. Navigate to the folder where you saved your files and Load **01 - Smart Zones Guide Samples.tif** as your template image.

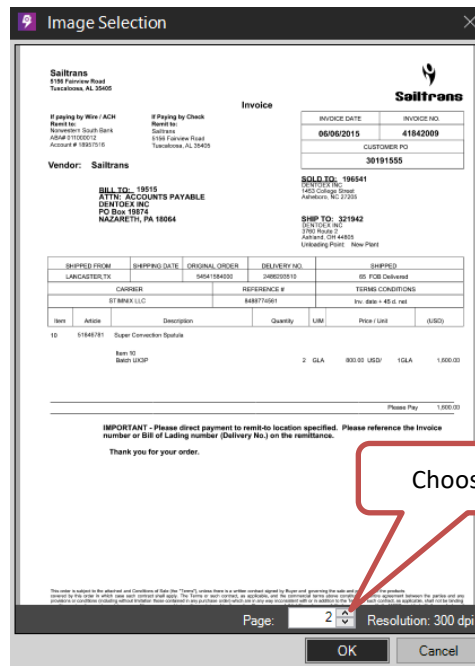


When prompted, select **Yes** to load only one page of the multi-page document as your template. Although you will eventually use each of the invoice variations to verify, you will only be working with one at a time. You will start with the invoice on **Page 2**.



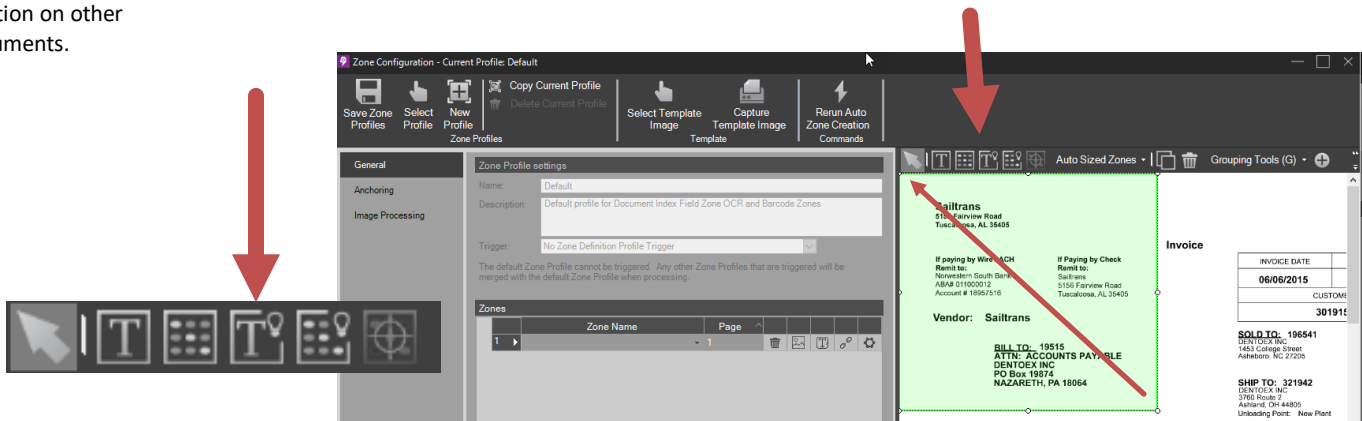
Select **Page 2**, the Sailtrans Invoice. You will be using this document to draw the initial zones.

## SMART ZONES QUICK START GUIDE



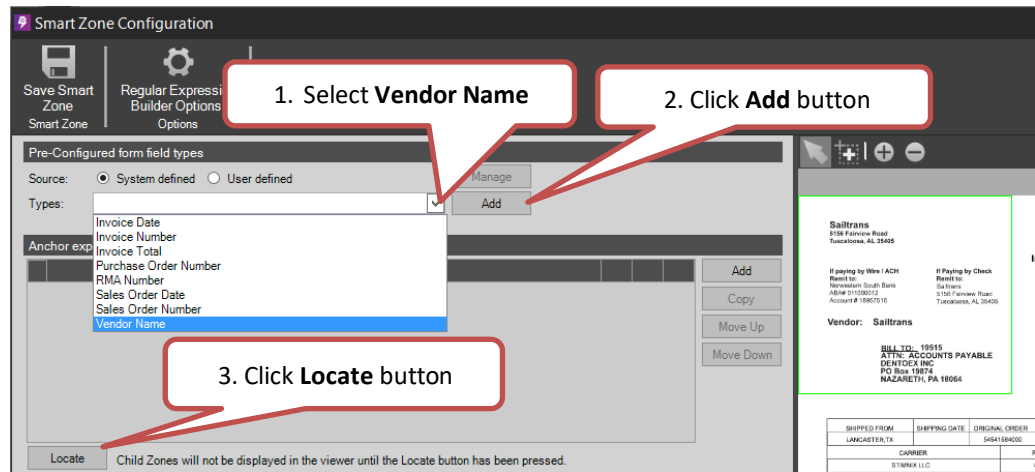
☆ Note:  
You are intentionally using an oversized Smart Zone in case the data you are seeking falls in a different position on other documents.

Select the Smart Zone tool and draw your first oversized zone from the bottom-right of the **Bill to** to the top-left of the **Page**. This zone will be large enough to capture the **Vendor Name** from each of the invoices.



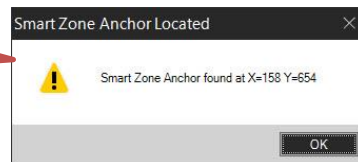
Upon releasing the mouse button, the Smart Zone Configuration window will appear. Select **Vendor Name** from the **Form Field Types** dropdown, then click **Add**.

★ **Note:**  
If you wanted to use an anchor that did not fall into one of the pre-defined Form Field Types, you could write a custom Regular Expression to match that anchor instead.



Click **Locate** - an alert notifies us that the anchor was found. Notice that the word **Vendor** is now highlighted in a green box. Based on the selection in the Form Field Types dropdown, PScapture automatically selected this label as the anchor.

Alert notifies us that the anchor was located – click **OK**

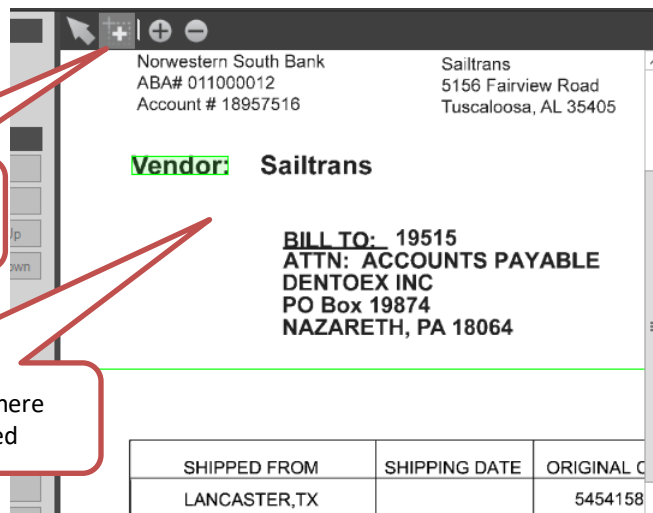


**Vendor:** Sailtrans

In order to enhance precision while drawing the child zone, you will zoom in on the anchor using the **Selection Zoom** tool.

Click the **Selection Zoom** tool bar button

Zoom in around the region where the Vendor anchor is located

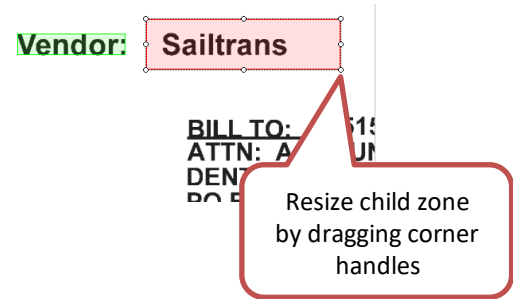
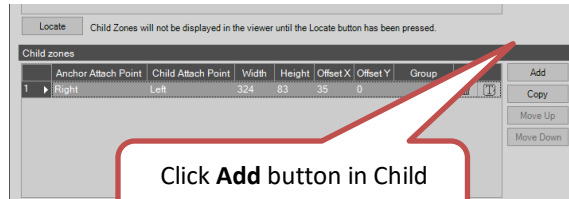


## SMART ZONES QUICK START GUIDE

### ☆ Note:

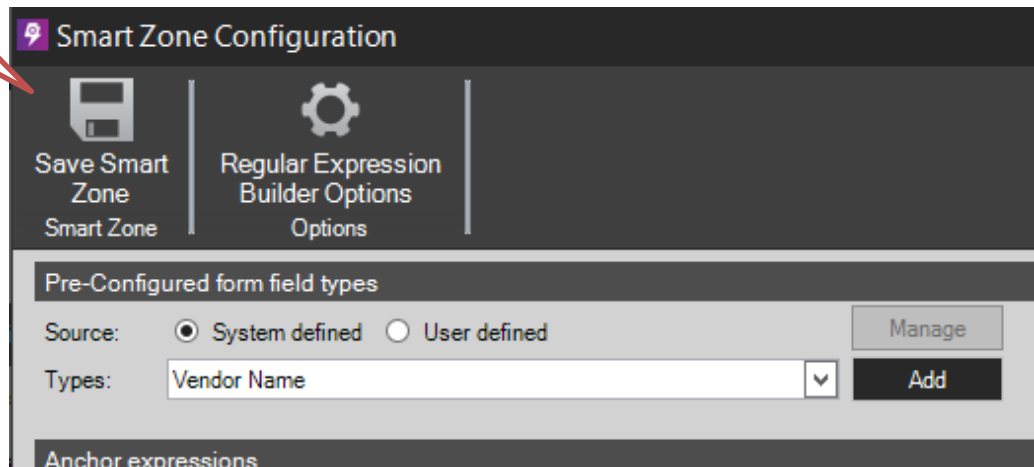
It's a best practice to leave some extra space around the data in the child zone but try not to overlap any extraneous text.

Next, you will add a child zone using the **Add** button in the lower middle of the Smart Zone Configuration window. Manually resize the child zone so that it overlaps the entire company name to the right of the word **Vendor**:



Click **Save** button

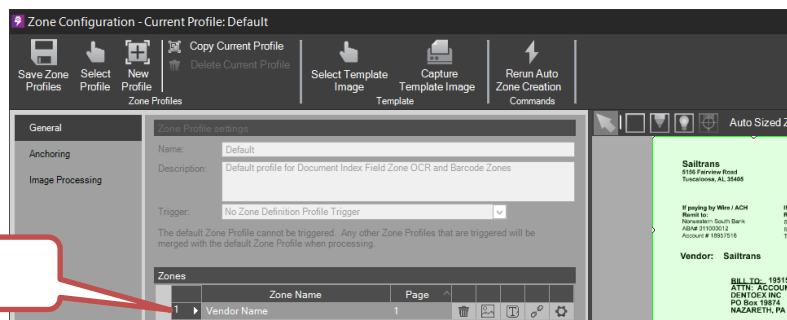
Now that you have finished building the first Smart Zone, click the Save button, from the upper-left corner of the Smart Zone Configuration window.



This will take you back to the **Zone Configuration** window.

Now you will assign a name to the newly created Smart Zone. Because this zone is designed to extract the name of the vendor, you will name the zone **Vendor Name**.

Name Zone



## SMART ZONES QUICK START GUIDE

Now repeat the process for the other 3 Smart Zones on the Sailtrans invoice. Position the other zones as shown here:

Vendor Name

**Sailtrans**  
5156 Fairview Road  
Tuscaloosa, AL 35405

**Paying by Wire / ACH**  
Remit to:  
Newmarket South Bank  
ABN 611000012  
Account # 10001210

**Paying by Check**  
Remit to:  
Sailtrans  
5156 Fairview Road  
Tuscaloosa, AL 35405

**Vendor: Sailtrans**

**ATTN: ACCOUNTS PAYABLE**  
**DENTOEK INC**  
**PO Box 19874**  
**NAZARETH, PA 18064**

Invoice

**Sailtrans**

INVOICE DATE	INVOICE NO.
06/06/2015	41842089
CARTON	
30191555	

**SOLD TO: 198541**  
**DENTOEK INC**  
1430 Cottage Street  
Aurora, NC 27205

**SHIP TO: 321942**  
**DENTOEK INC**  
1753 Route 2  
Aurora, OH 44202  
Unloading Point: New Plant

**2 Overlapping Zones - Invoice Number and Invoice Date zones overlap exactly**

☆ Note:

In order to determine the proper size and placement of the Smart Zones, the creator of this guide looked through all the sample documents in advance. When designing your own projects, this is something you would need to do on your own.

SHIPPED FROM	SHIPPING DATE	ORIGINAL ORDER	DELIVERY NO.	SHIPPER		
LANCASTER, TX		5454-154000	240620310	BS FOG Delivered		
CARRIER	REFERENCE #	TERMS CONDITIONS				
STIMEX LLC	8886774561	Inv. date + 45 d. net				
Item	Article	Description	Quantity	Unit	Price / Unit	(USD)
13	51845701	Super Conversion Spindle				
	Item 10	Batch LUGP	2	BLA	680.00 USD	1384.00
						Placeholder 1.800.00

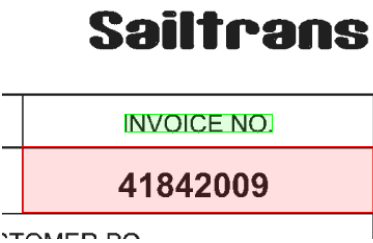
**IMPORTANT - Please direct payment to remit-to location specified. Please reference the invoice number or Bill of Lading number (Delivery No.) on the remittance.**

**Thank you for your order.**

**Invoice Total**

1. Draw the Smart Zone.
2. Select the proper Form Field Type and click Add.
3. Click the Locate button to find the anchor on the template image.
4. Use Selection Zoom around the anchor to enhance precision.
5. Click the lower Add button to create a child zone.
6. Re-position and re-size the child zone appropriately.
7. Save out of the Smart Zone Configuration window.
8. Name the zone according to the data it will extract.

If necessary, position the child zone below the anchor instead of to the right:



Four (4) Smart Zones have been created with your typed descriptive zone names.

Zones						
	Zone Name	Page ^				
1	Vendor Name	1				
2	Invoice Number	1				
3	Invoice Date	1				
4	Invoice Total	1				

Now that all 4 Smart Zones have been configured correctly on the Sailtrans invoice, you will replace the Sailtrans template image with a sample invoice from the next vendor, Voltquote.

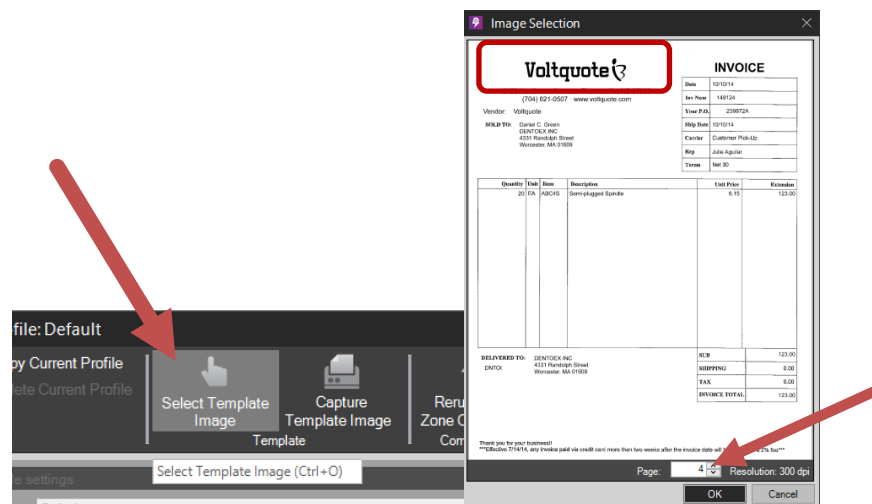
You will verify that each anchor falls within the Smart Zone on this Voltquote document as well. Then you will open the individual Smart Zone Configurations to ensure that the child zone you drew for Sailtrans will also work for Voltquote.

When the required data *does not* fall within the child zone, you will rectify this by either creating a second child zone or enlarging the original child zone.

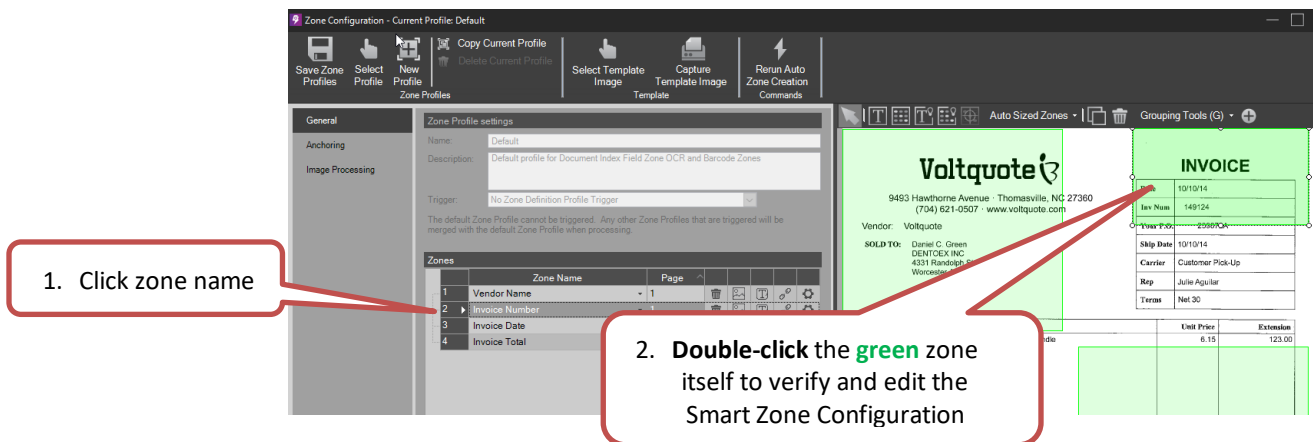
☆ Note:  
Re-loading the template image will only replace the image of the page. The zones that you created will remain intact.

Click the **Select Template Image** button, and replace the Sailtrans Template Image, which was page 2, with the Voltquote Invoice, which is Page 4 from the sample document, **01 - Smart Zones Guide Samples.tif**.

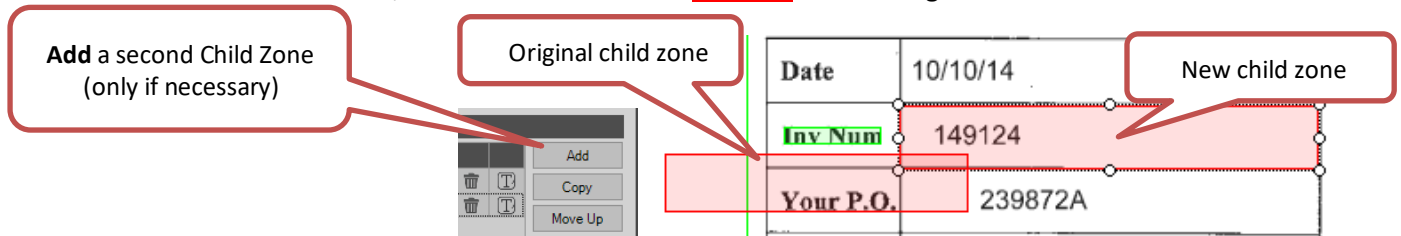
## SMART ZONES QUICK START GUIDE



With the Voltquote document loaded, select the zone that you want to verify by clicking its name on the **Zones** tab. Then double-click the corresponding **green** zone on the template image.

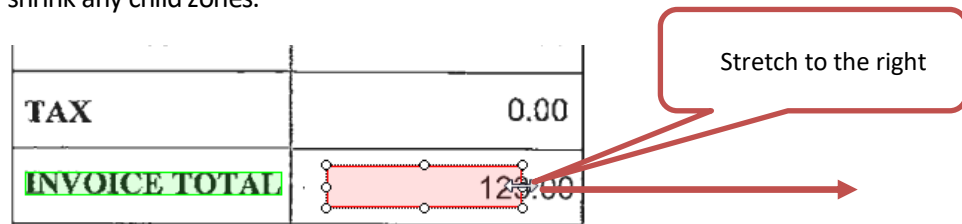


When the value is found in a different direction from the anchor than on the Sailtrans invoice, **Add** a second child zone. **DO NOT** alter the original child zone.



☆ Note:  
In order to determine which of the two child zones to use, PSICapture will look for a certain pattern of characters, which you will specify using a Regular Expression.

If the child zone is in the correct position, but not large enough to enclose the full value, use the zone handles to expand the child zone so that it contains the entire value. **DO NOT** shrink any child zones.



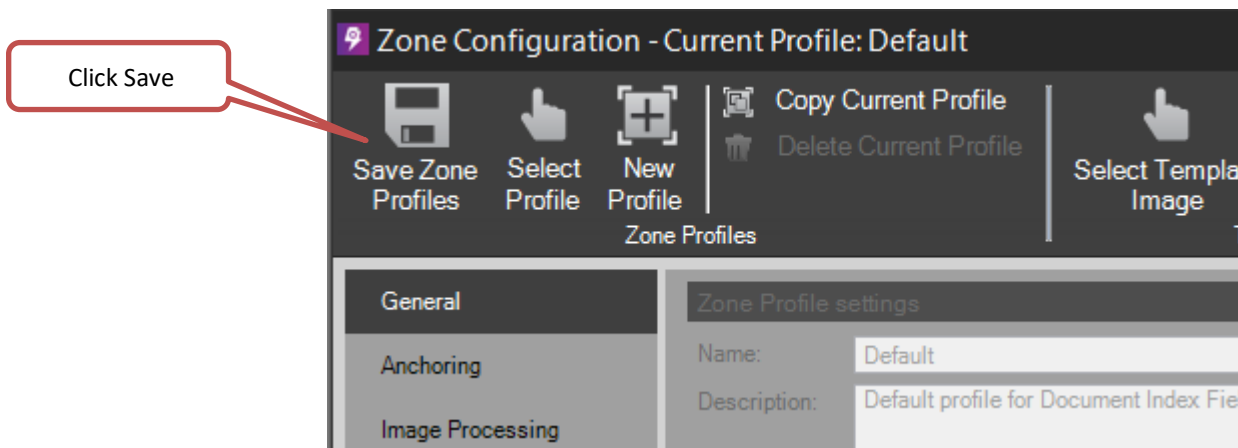
Repeat this process for all 4 fields on the Voltquote Invoice, ensuring that the anchors are successfully located, and that one child zone completely encloses the index data that you intend to extract.

Once Voltquote is complete, overwrite the template image again by loading:  
**Page 1 of 01 - Smart Zones Guide Samples.tif**, which is the TamTech invoice.

☆ Note:  
When Smart Zones overlap, remember to select the zone name that you would like to edit from the list prior to double-clicking.

1. Verify that intended anchors always fall within the green zones.
2. Open each of the 4 Smart Zones by double-clicking them.
3. Verify that one of the child zones completely encloses the intended data.
4. Either stretch the original child zone or create another child zone to fully enclose the data as necessary. Do not shrink any child zones and do not move the zones.

Now that the Zone Configuration has been refined in such a way that it should extract the required index values from all 3 vendors' documents, you will **Save** the Zone Configuration.



This returns you to the Indexing-Document Fields Tab.

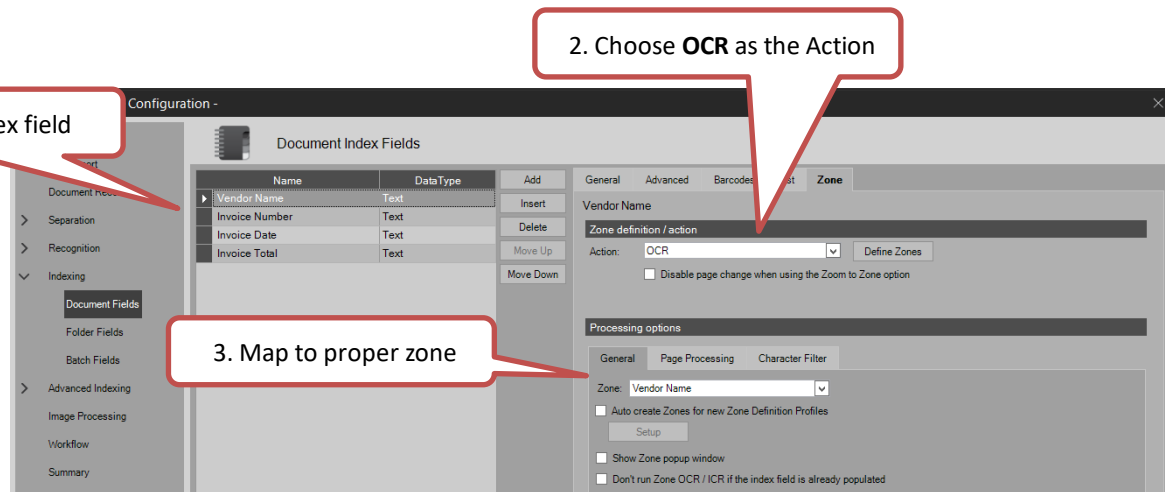


For each index field, go to the **Zone** tab and select the **Action** as **OCR** from the dropdown. You choose this option because that is the type of recognition engine that you will be using to read the zones.

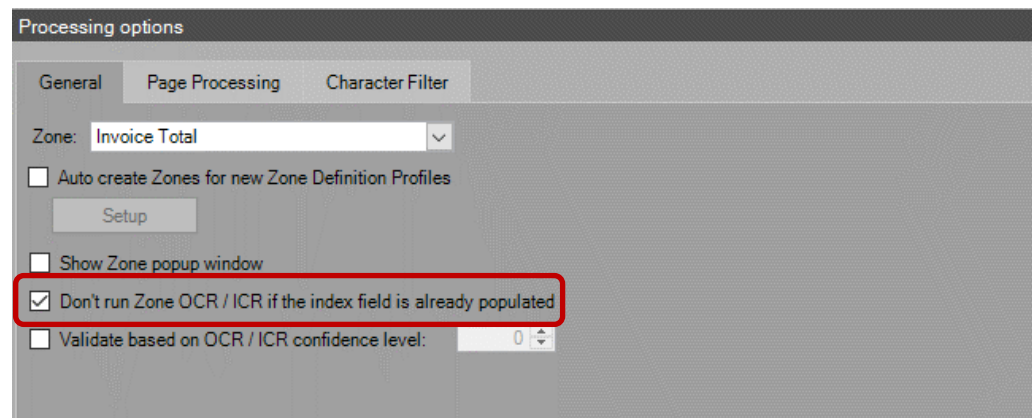
Then you will ensure that each Index Field is mapped to the proper zone.

☆ Note:  
Index Fields will be automatically mapped to the proper zones if their names match **exactly**.

☆ Note:  
Don't forget to repeat these steps on the Zone tab a total of 4 times (once for each Index Field).



For each Index Field, at the bottom of the **Zone** tab, check the box that says **Don't Run Zone OCR if field is already populated**. This is very important when using multi-child Smart Zones. It means that once the software finds a value in one child zone that matches the defined regular expression pattern, it will extract that data and stop looking in any other child zones.

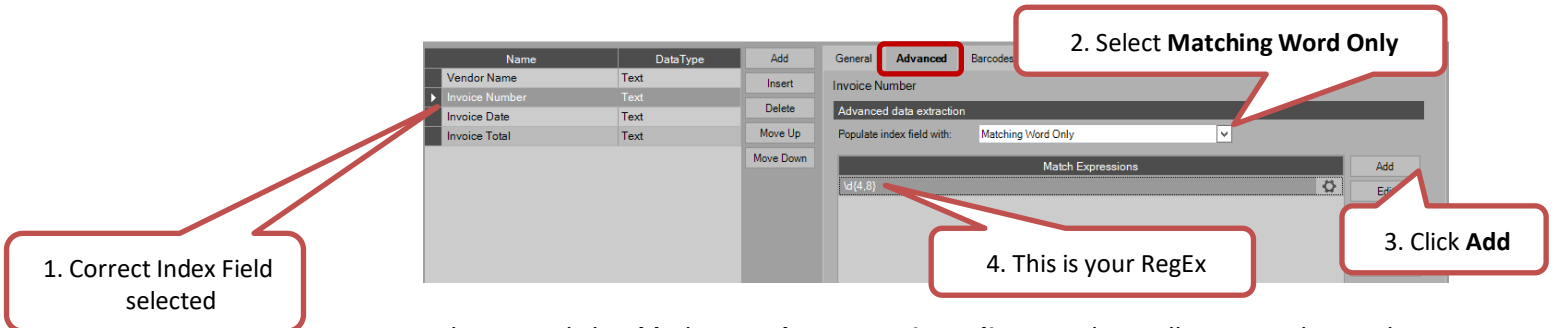


Next you will add Regular Expressions (RegEx) to the **Advanced** tab in Indexing. A Regular Expression is a sequence of characters that forms a search pattern. Each index field is now associated with a certain Smart Zone. You will configure the software to look for specific patterns of text within that Smart Zone's child zones.

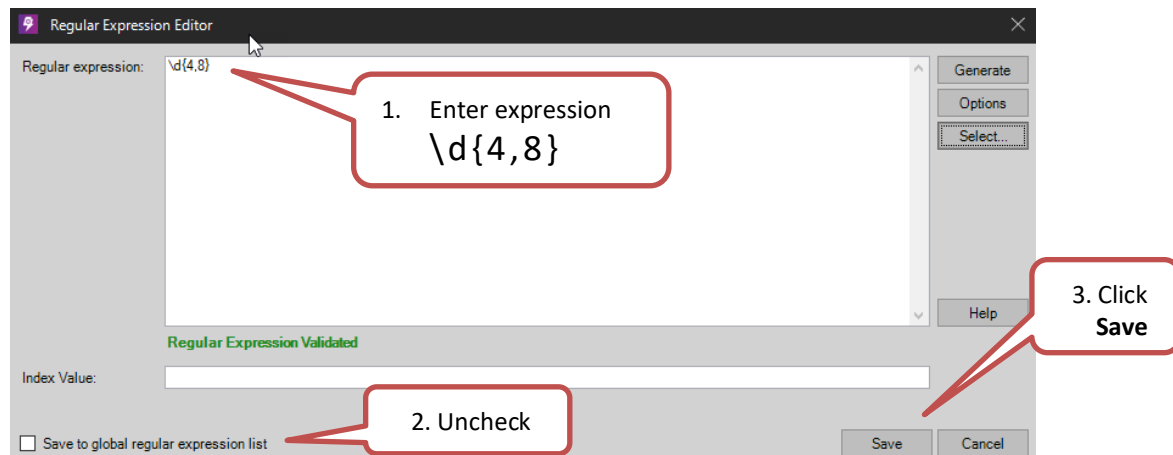
☆ Note:  
If you forget to check  
**Don't run Zone OCR if  
field is already populated**,  
the software might  
completely overlook some  
successful RegEx matches.

In reviewing the invoices, you know that all of the Invoice Numbers are between 4 and 8 digits. For the **Invoice Number** index field, you will add a RegEx to the **Advanced** tab to look for such a pattern. If the OCR engine finds *anything other than* this pattern in the first child zone it evaluates (such as a string of alphabetical characters), it will move on to the next child zone to continue looking for the correct pattern.

The **Vendor Name** Field does not require a Regular Expression, so you will start with the **Invoice Number** Field.



When you click **Add**, the **Regular Expression Editor** window will appear. This is where you type your RegEx.



☆ Note:  
The **global Regular  
Expression list** is a place to  
save your RegEx for use in  
other Capture Profiles. You  
will not need to use this  
feature here.

You will be entering Regular Expressions for 3 of the 4 fields (Invoice Number, Invoice Date, and Invoice Total). From the **Advanced** tab, for each field:

1. Select the Index Field Name on the left.
2. Select **Matching Word Only** from the dropdown.
3. Click **Add**.
4. Enter your **Regular Expression**, as specified here:

☆ Note:

If typing your RegEx, be sure to correctly differentiate between:

Curly braces { } and  
Square brackets [ ]

### Invoice Number

**\d{4,8}**

A series of between 4 and 8 digits.

### Invoice Date

Select from Global List

Regular Expression Editor

Regular expression:

Regular Expression Validated

Index Value:

☐ Save to global regular expression list

Generate Options Select... Help Save Cancel

Click Date

Regular Expressions Manager

Type	Description	Expression
System	Date	([012][01-9])(?<[0-9][0-9][0-9][0-9])([12]/[0-9])
System	Email Address	([0-9a-zA-Z]([-.\w])*[0-9a-zA-Z])*@[0-9a-zA-Z]
System	Phone Number	[01]?[- ]?(?([2-9][0-9][2-9])([2-9][0-9][2-9])[- ]?([0-9])+\$
System	Positive Integer	[0-9]+
System	Social Security Number	[0-9](3)-[0-9](2)-[0-9](4)
System	Time	((0?[1-9] 1012)([0-5]d)(0.2)([17](AaPp Mm))
System	Weekday	(Sun Mon (T ues hurs thu f?) (Fri)(day(\. )? v)
System	Zip Code	[0-9](5)([- /]?[0-9](4))?

Add Edit Delete Select Cancel

Click Select and Save

### Invoice Total

**[0-9,.]+**

Simple re-usable  
Currency RegEx

Any uninterrupted series of digits, decimals, and commas.

Once you have completed the above steps on both the **Zone** tab and the **Advanced** tab for all the Index fields except Vendor Name, you have finished setting up **Indexing**.

## Advanced Indexing Tab

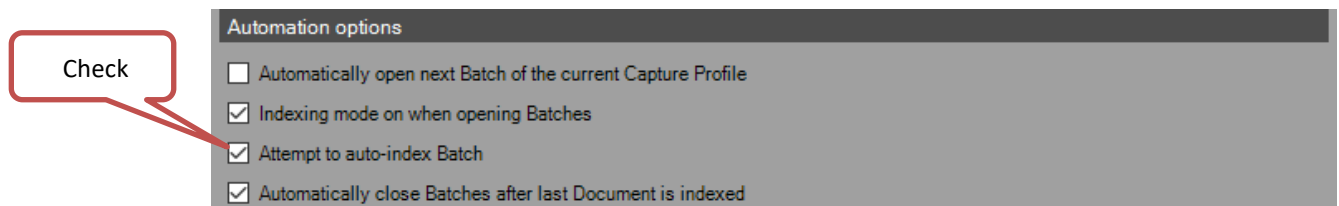
Since no changes are required here, you can skip this tab.

## Image Processing Tab

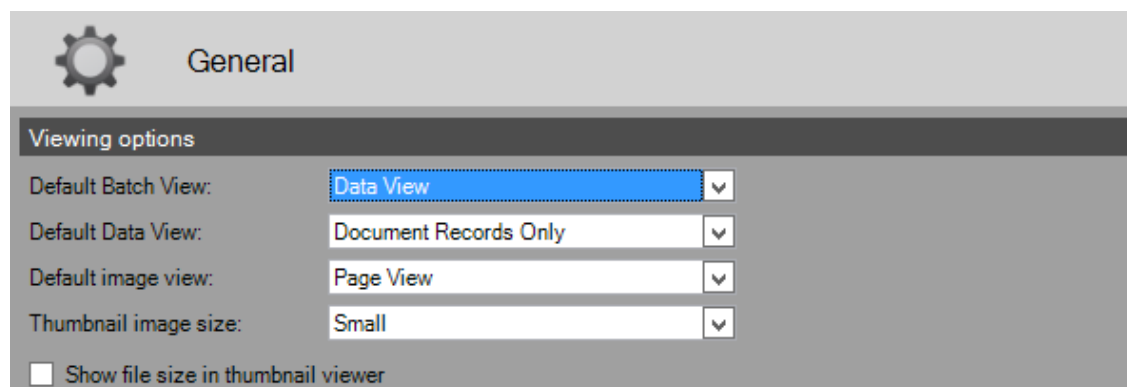
Since no changes are required here, you can skip this tab as well.

## Workflow Tab

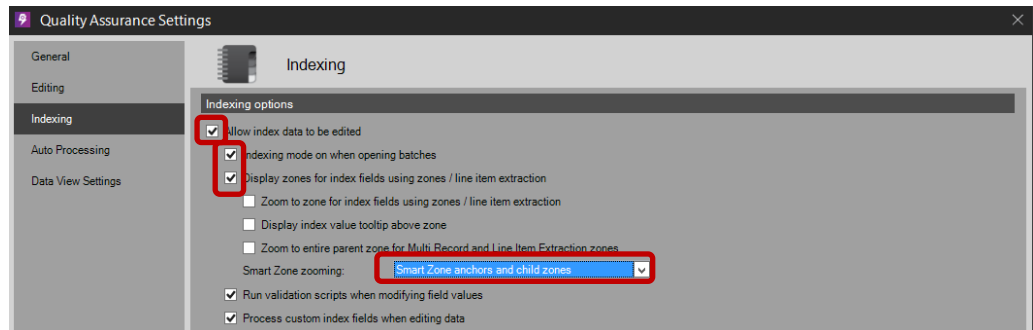
In this tab you will build the PSIsapture workflow. Select **Index** from the left and click **Add**. You will want to make only one change – check the **Attempt to Auto-Index Batch** checkbox from the top in the Automation section. This will auto-index each document without requiring user interaction from the Index Module. Click **Save**.



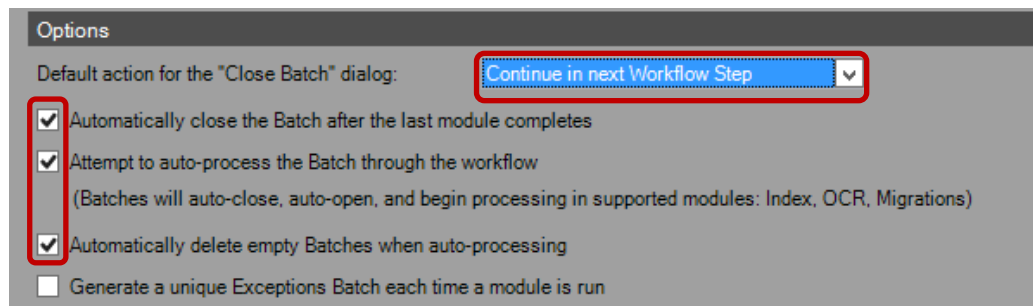
Next you will select **Quality Assurance** from the left and click **Add**. Change the Default Batch View from Tree View to **Data View**. This will cause the values you will have extracted during Indexing to be displayed when you enter the Quality Assurance workflow module.



On the Quality Assurance – Indexing sub tab, check the 3 checkboxes outlined in the below screenshot and choose **Smart Zone anchors and child zones** from the dropdown. This will display the exact location from which each index value was extracted as you verify the data from the Quality Assurance screen. Click **Save**.

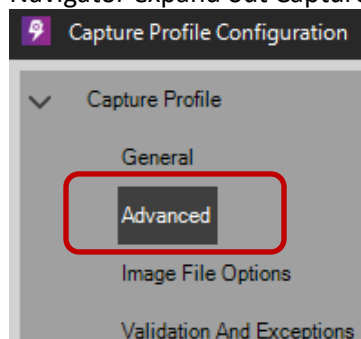


Next, in the Workflow options (at the bottom of the tab), check the top 3 checkboxes and change the dropdown to match below. This will automate the workflow, so that the Capture Profile will process the documents without requiring user interaction after the Capture/Import module.

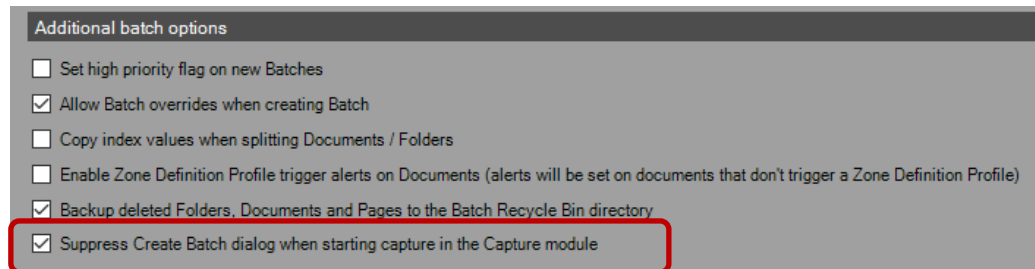


## Suppress Create Batch Screen

When running a batch through PSICapture the Create Batch screen is displayed by default. We will not have a need to change any settings on that screen in this profile, so we are going to suppress that screen from being displayed. On the left-hand Navigator expand out Capture Profile and select the Advanced tab.

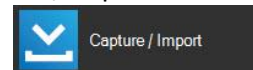


In the Additional batch options section, we want to check the 6<sup>th</sup> checkbox titled “Suppress Create Batch dialog when starting capture in the Capture module”

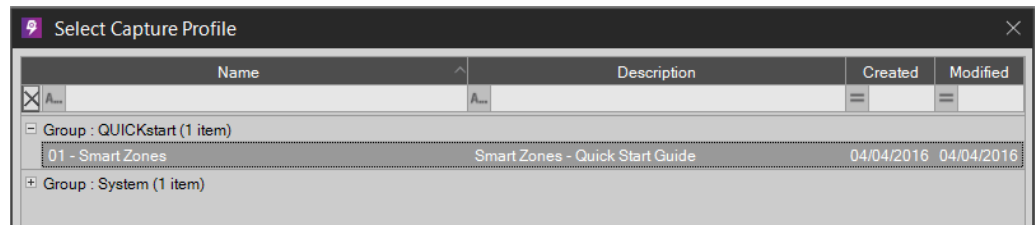


## Run the Capture Profile

You can now click **Save** in the lower right-hand corner. Click **Save** again from the Configuration screen and run the Capture Profile by clicking Capture / Import tab.

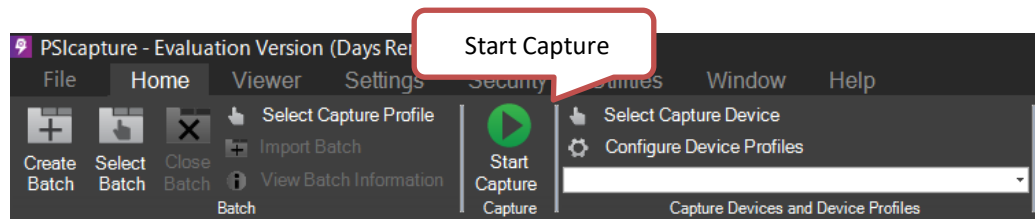


Then select your new 01 – Smart Zones capture profile.

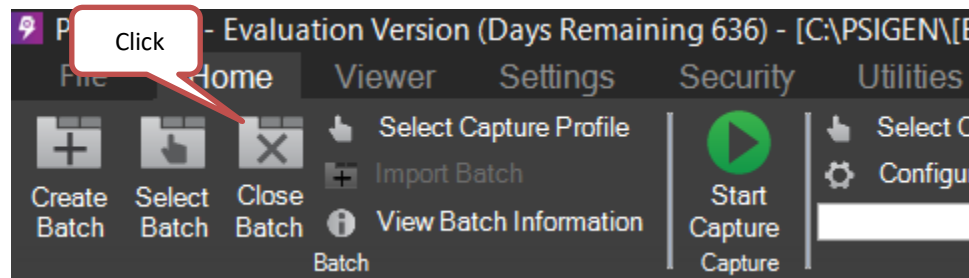


Using the File Browser pane, navigate to where you saved the files for this guide. You will select the folder on top and then select **01 – Smart Zones Guide Samples.tif** from the lower pane. Highlight the file and select the green **Start Capture** button (or Control+S) to begin importing the documents into PSICapture.

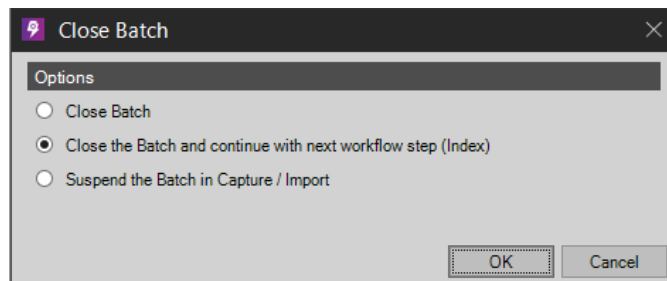
☆ Note:  
If you do not see the File Browser toward the left of the PSICapture screen, click **Select Capture Device...** icon to switch to the **Import** mode.



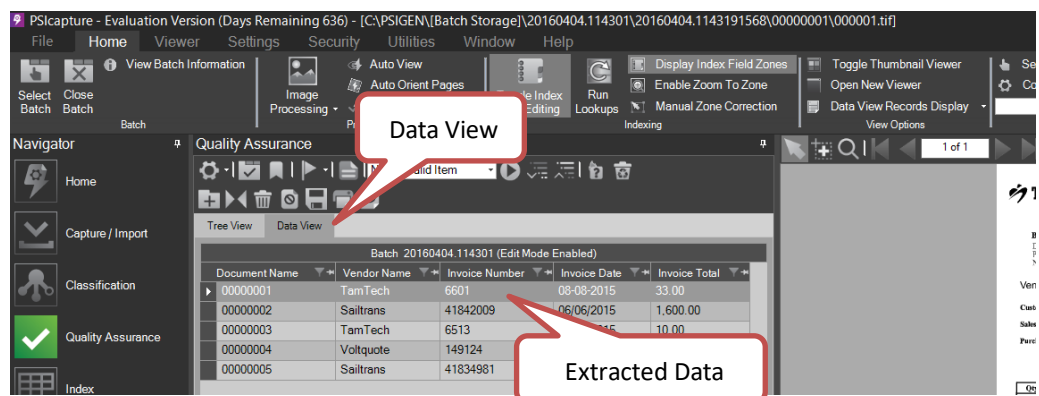
When the file has finished importing, you will click the **Close Batch** button.



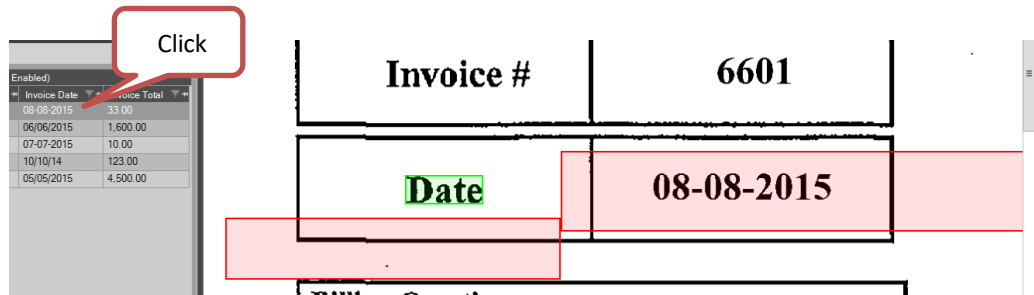
Ensure the selection for **Close Batch and Continue in next Workflow step (Index)** is selected. Click **OK**.



The Index step will attempt to auto-index the invoices and will then pass the batch directly to the Quality Assurance step, from which you will be able to review the extracted data.



Click inside each individual cell of the Data View to view the exact zone on the page from which the data was extracted. Even though the anchors float around the page and sometimes there are multiple child zones, PSICapture knows exactly where to find what it's looking for in every case.



CONGRATULATIONS!!  
You've just configured and tested a  
Capture Profile using Smart Zones.