

## Configuring

[Configuring](#) / [Configuring document processing features](#) / [Using RICOH ProcessDirector Plug-in for Adobe Acrobat](#) / [Working with document properties](#)  
/ Defining a document property

## Defining a document property

Document processing features can store document property values in the RICOH ProcessDirector database. The features rely upon document properties for later downstream processing of PDF files in RICOH ProcessDirector.

### ⚡ Note:

- Read the document property overview section to ensure you understand how document properties are used in RICOH ProcessDirector so you can take full advantage of your document processing features.
- Before you begin this procedure, verify that the most recent RICOH ProcessDirectordefinitions.zip file is loaded.

To define a document property:

1. Open a PDF file in Adobe Acrobat Professional and either load a control file that contains a page group definition or define a page group.
2. Left-click just above the top left corner of the data that you want to capture. Drag the mouse to draw a box around the data. You can later view the extracted values to verify your selection.

### ⚡ Note:

- The data to capture can be text or DataMatrix barcodes encoded as images.
- Make the box big enough to capture the longest occurrence of the data in your PDF files. Some characters in a PDF file have a larger white space buffer than other characters. For example, the left edge of a large uppercase letter might have up to a tenth of an inch of white space buffer that you need to select to capture that letter.

3. Select **Define Document Property** from the menu.
4. Select a RICOH ProcessDirector document property from the list or type a document property name into the field. Do not use any special characters (such as @, #, \$, %, or - (dash)) or spaces, because the RICOH ProcessDirector **IdentifyPDFDocuments** step might fail. You can use periods and underscores.



### ⚡ Note:

- If you do not see the RICOH ProcessDirector document property that you want to use, reload the definitions.zip file.
- You can define a document property more than once.  
For example, if text in your PDF file is variable, you might need to extract the postal code from different locations. You can define your postal code document property twice, as long as you define different conditional placement rules that specify the pages from which the property is extracted.  
If you define the same document property in two different ways in the document, and each of their conditions are met, the last value extracted is used.

5. Define which type of data to extract values from.
  - If you selected an area that only contains text, select **Text** under **Select from**.
  - If you selected an area that only contains barcodes, select **Barcode image** under **Select from**.
  - If you selected an area that contains both text and barcodes, select both **Text** and **Barcode image**.  
The text data is placed before the barcode data in the extracted string without an indicator of where the text data ends and the barcode data begins.

### ⚡ Note:

- We recommend using black barcodes. Using other colored barcodes might have unpredictable results.

6. Specify the page in each document from which document property data will be extracted. Do either of these:
  - Select **Pages based on a rule**, and then select a rule from the drop-down list. The default rule is **First Front Only**. You can also:
    - Click the  icon to define a new rule.
    - Click the  icon to go to the Rules Manager.

### ★ Important:

- The **Last Back**, **Last Front**, and **Last Page** rules do not work with the extraction of document property data.


- Select **Specific pages** and type the page in each document that you want.

If you specify multiple pages, RICOH ProcessDirector Plug-in for Adobe Acrobat extracts the document property data from the last specified page in each document. Examples:

- You specify pages 2–4. If a document has four or more pages, the document property data is extracted from page 4. If a document has three pages, the document property data is extracted from page 3. If a document has two pages, the document property data is extracted from page 2.
- You specify pages 2, 4. If a document has four or more pages, the document property data is extracted from page 4. If a document has 2–3 pages, the document property data is extracted from page 2.
- You specify pages 2–n. Because n represents the last page, the document property data is extracted from the last page if the document has two or more pages.

★ **Important:**

- If you specify only page n, RICOH ProcessDirector Plug-in for Adobe Acrobat does not extract the document property data from any page in a document.



7. **Optional:** Select the edit icon  to display a Modify Text window where you define one or more modifier extraction rules to extract the exact document property you need.

1. Choose one of the following modifiers:  
Content modifiers

Modifier	Action
Remove Character	<p>Type one character or a blank character (use the space bar to type a blank character) that you want to remove from the value. The character is case-sensitive. Then select one of these buttons:</p> <ul style="list-style-type: none"> <li>■ <b>Remove all instances of the character</b> The specified character is removed from all positions in the value. For example, an account number is: 324-1443255-11. You can type a - to remove all - characters from the value, producing 324144325511.</li> <li>■ <b>Remove leading characters</b> The specified character is removed from the beginning of the value. For example, if you type a blank character, all blank characters are removed from the beginning of the value.</li> <li>■ <b>Remove trailing characters</b> The specified character is removed from the end of the value. For example, if you type a blank character, all blank characters are removed from the end of the value.</li> <li>■ <b>Remove leading and trailing characters</b> The specified character is removed from the beginning and end of the value. For example, if you type a blank character, all blank characters are removed from the beginning and end of the value.</li> </ul>
Substring by Position	<p>Select <b>Beginning of Line</b> or <b>End of Line</b> from the <b>Starting From</b> list. Select a number for <b>First Position</b> to indicate the location of the first character in the text value. Select a number for <b>Number to Retain</b> to indicate how many characters are retained.</p>
Substring by Delimiter	<p>Type a character or a blank character in the <b>Delimiter</b> field to indicate where the text value is split into separate string segments. The character and the text string are case-sensitive. Select <b>Beginning of Line</b> or <b>End of Line</b> from the <b>Starting From</b> drop-down menu. Select a number for <b>First Position</b> to define the position of the delimiter in the text string. Select a number for <b>Number to Retain</b> to define the number of text string segments to retain.</p> <p>These examples show how to select text string segments by specifying a delimiter:</p> <ul style="list-style-type: none"> <li>■ For the account number 324-1443255-11, you can use - as the delimiter to split the value into these three text strings: 324, 1443255, and 11. Select <b>Beginning of Line</b>. To select the second and third text strings (1443255 and 11), select <b>2</b> for both <b>First Position</b> and <b>Number to Retain</b>.</li> <li>■ For the mailing address Eldorado Springs CO 80025, you can use a blank character as the delimiter to split the value into these four text strings: Eldorado, Springs, CO, and 80025. Select <b>End of Line</b>. <ul style="list-style-type: none"> <li>■ To select the zip code, select <b>1</b> for both <b>First Position</b> and <b>Number to Retain</b>.</li> <li>■ To select the state, select <b>2</b> for <b>First Position</b> and <b>1</b> for <b>Number to Retain</b>.</li> <li>■ To select the city, select <b>3</b> for <b>First Position</b> and <b>10</b> for <b>Number to Retain</b>. By specifying <b>10</b> for <b>Number to Retain</b>, you can select city names with up to ten words.</li> </ul> </li> </ul>
Pad with Character	<p>Select <b>Beginning of Line</b> or <b>End of Line</b> from the <b>Padding Location</b> list. Enter a character or a blank character as the padding character into the <b>Character to Pad with</b> field. Enter a number in the <b>Minimum Padded Text Length</b> field to define the minimum length of the text string. If the number of characters in the text string is less than this minimum length, padding characters are added until the text string equals the minimum length.</p>

When you use a modifier to define a text extraction rule, the **Text to Modify** field at the top of the Modify Text window contains the selected line plus any edits you make to the line. The **Modified Value** field to the right of a modifier displays the text that results when that modifier is applied to the text it received from either the modifier above it or the **Text to Modify** field (if you are defining the first

modifier).

2. Continue to apply modifiers until you extract the value you want from the selected line. Click the  icon to add a new modifier. The **Final Text** field below the list of modifiers displays the final modified value, after all modifier extraction rules are applied. For the selected modifier, the **Modifier Initial Text** field at the bottom of the window displays the value before the modifier is applied. The **Modified Text** field displays the value after the modifier is applied.
3. Use the modifier management icons near the top of the window to delete and reorder the modifier extraction rules. Use the  icon to delete the selected modifier extraction rules. Use the up and down arrow icons to reorder the rules. The rules are applied to the line in order from top to bottom.
4. Click the **OK** button to save the line extraction rule.
8. Click **OK** to create the document property.
9. Click **Ricoh ⇒ View Document Property Values** and scroll through several documents in your PDF file to verify that RICOH ProcessDirector Plug-in for Adobe Acrobat is extracting the correct document property values for each document.
10. When you are ready to save all your enhancements to the PDF file, including the new document property definition, click **Ricoh ⇒ Save Control File**.
11. In the RICOH ProcessDirector **IdentifyPDFDocuments** step, specify the name and location of the control file that contains the document property definition.

Parent topic: [Working with document properties](#)