Best Practices for Scanning in Laserfiche 8

White Paper

October 2008

Laserfiche®

1

Table of Contents

Scanning Strategy

Laserfiche Scanning helps you quickly and easily convert paper documents to images for storage in your electronic repository. The Laserfiche Scanning interface provides a wide variety of options for processing documents are they are scanned, to make them easier to find, easier to read, and generally more informative, A good scanning procedure is key to creating useful, well-organized repositories, and to keeping the scanning process fast and accurate.

When creating a scan procedure, you should keep several elements in mind. Determine how you will capture your images and what scanning engine you will use. Set procedures for parts of the process specific to Laserfiche Scanning, including assigning metadata to documents; processing images to improve aesthetics, readability, and the Optical Character Recognition (OCR) process; and sending the documents to Laserfiche.

Organizing Documents and Batches

If you are scanning a large batch of documents, you will want to make sure the individual documents within the batch can be quickly and efficiently identified so they can be sorted, assigned metadata, and stored appropriately in Laserfiche. Creating consistent organization procedures before scanning can greatly enhance the speed and accuracy of the scanning process.

Slipsheets help scanner operators visually distinguish between different documents. Also make sure each document has all the necessary information in a place that is easy to locate and read, such as a cover sheet. Laserfiche Scanning offers features to remove slip sheets and cover sheets before the documents are stored in your repository.

Choosing Your Scan Engine

There are several possible options for interfacing with your particular scanner. Your organization may choose to use one or several, depending on your scanning procedure, available equipment, and preferred image processing methods.

Kofax

Laserfiche Scanning can communicate with any scanners that use Kofax Adrenaline cards and software to capture images. Kofax will only be installed if your Kofax Adrenaline Capture Engine (ACE) or Virtual ReScan (VRS) has been installed. These must be purchased separately. Kofax software should be installed prior to installing Laserfiche Scanning. If this is not possible, repair the Laserfiche Scanning installation after installing the Kofax software.

ScanConnect/ISIS

ScanConnect enables you to use ISIS scanning drivers. ScanConnect must be purchased separately. A collection of ISIS drivers is included with ScanConnect. A list of supported scanners can be found on the [Laserfiche Support Site](#). If your scanner is not listed, manually install any ISIS drivers your scanner's manufacturer has provided.

Note: ScanConnect cannot be used in conjunction with the Kofax ACE CD.

TWAIN

TWAIN is automatically installed with Laserfiche Scanning, but a TWAIN driver appropriate for the desired scanner must also be installed on your machine for you to be able to use this scan source. This driver will be provided by your scanner's manufacturer if the scanner supports TWAIN.

For any given scanner, TWAIN is usually slower than ISIS, so if speed is a concern and ISIS is available you may prefer to use ISIS.

Universal Capture

Universal Capture is automatically installed with Laserfiche Scanning and works as a virtual scanner by capturing images already stored on your computer or network drive. This enables you to:

- Use any scanner and scanning software you like to scan new images while still using Laserfiche Scanning features.

- Apply Laserfiche Scanning features to images you have scanned in the past.

Universal Capture can scan TIFF, bitmap, and JPEG images.

Basic and Standard Modes

When you open Laserfiche Scanning, the first choice you make is between basic and standard scanning modes, unless you have turned off the Laserfiche Scanning Mode dialog box in your Laserfiche Client. To turn it and other hidden dialogs back on, go to **Tools**, **Options**, **Prompts** and select **Scan Mode**.

Basic Mode requires minimal configuration. It provides a few basic enhancements and the option to set default document properties or fill in properties for each document. It is ideal for beginning and intermediate users with relatively simple scanning needs, or for short scanning jobs by any user.

4
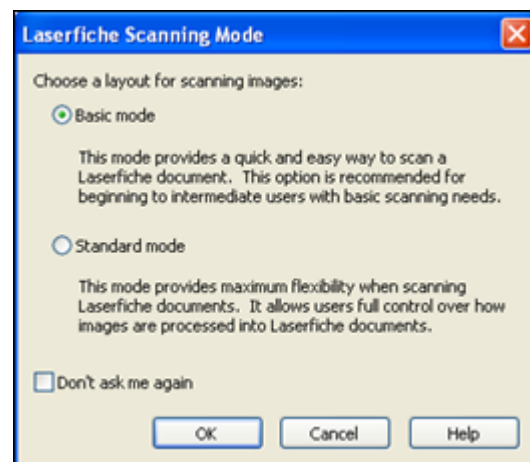
Standard Mode provides much more flexibility, including a broad variety of enhancements and properties to give users full control over images. Standard mode is best for more advanced users.

Streamlining Metadata

When you set up templates for scanned documents, don't include too many required fields. Since a scanner operator cannot file a document until all the required fields have been filled in, required fields can slow down the scanning process considerably, particularly if the required information is missing or difficult to find.

Enhancement Best Practices

Scanned documents may not always be ideally formatted for electronic reading. They can be smudged or discolored, have margins that are too large or too small, or contain blank pages and irrelevant information. In order to fix these potential defects and thereby improve the results of Optical Character Recognition (OCR), Laserfiche Scanning allows users to perform enhancements on scanned documents.

The Laserfiche Scanning help file describes the function of each image enhancement and processing operation. However, the order in which these operations are applied can have a significant effect on the end results, as well as on the time it takes to process scanned images. In Standard Scanning, you can choose the order in which to apply them.

The following list describes the most effective order to perform image enhancements and processes. It is not necessary, and in most situations not desirable, to perform every enhancement listed here; the list suggests the preferred order for those you do perform. For instance, if you want to despeckle, crop, and deskew your image, you would use this reference to see that you should deskew first, crop next, and despeckle last.

Sometimes a different order might work better for your documents. Experiment with different combinations to see what works best.

1. **Resize:** This enhancement will always reduce your image quality, so it should be used very sparingly. If you need to change the image size, it is better to do so by scanning at a lower or higher resolution.

2. **Remove Pages:** Removing unnecessary or blank pages before performing any other enhancement or process ensures that the scanning program will not waste time processing pages that will eventually be thrown away. Remember that deleting pages is permanent, so be sure to copy any information you need from cover sheets before deleting them.
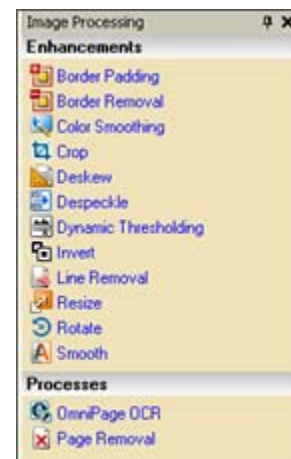
3. **Rotate** and **Deskew:** Rotate will turn an image upright by specified 90 degree increments, and Deskew automatically straightens crooked images. Both improve the accuracy of OCR, which expects characters to be right side up.

4. **Crop** and **Border Removal:** Cropping unnecessary regions can save processing time when you apply other enhancements. Because they reduce the overall size of the image, they also save disk space when the document is stored in Laserfiche.

   The exception is the Color Smoothing enhancement. Color Smoothing requires a sufficiently large region of blank background to serve as a "reference" region. If your cropping operation will not leave such a sample region, perform Color Smoothing before you crop.

5. **Color Smoothing:** This enhancement averages and evens out the background color of the image. Color Smoothing can only be applied to color images and improves the results of Dynamic Thresholding, so it should be performed first.

6. **Dynamic Thresholding (Color Removal):** Dynamic Thresholding converts color or grayscale images to black and white. This can greatly improve the performance of OCR processing.

7. **Invert:** If your image has white text on a black background, use the Invert function to change it to black text on a white background. This enhancement will work on color or grayscale images, but it will probably have the best results on black and white.

8. **Border Padding:** This enhancement adds extra space around the edges of the image. This allows more room for stamps or other annotations added later. The extra space will always be white, so if you use it on a dark image and then invert the colors, you will end up with a dark border on a light image. To keep the border color consistent with the rest of the background, use Border Padding after Invert.

9. **Line Removal:** This enhancement can remove lines that run through text and interfere with OCR. It will work on color as well as black and white images, but will be more efficient on black and white.

10. **Despeckle:** Despeckle only works on black and white images, and it expects images that have black text on a white background. Therefore, it is best to run it after Dynamic Thresholding and Invert.

11. **Smooth:** Use this enhancement to even out the shapes of your text. Like Despeckle, this enhancement expects black text on a white background and works on black and white images. Running Smooth after Despeckle will slightly improve processing speed, since Despeckle will already have removed some of the extraneous marks around the text.

12. **OCR:** OCR last. Most image enhancements are specifically intended to make text more readable both for users and for the OCR engine, so you should always make OCR your last step.

Laserfiche®

Best Practices for Scanning in Laserfiche 8

October 2008

Author: Regina Carns and Snow Tempest

Contributing Author: Ed Heaney

Editor: Constance Anderson

Technical Editor: Miruna Babatie