

productinfo

RECOGNITION

Index Automatically Using Barcode and Text Recognition

RECOGNITION gives you the power to file documents in DocuWare even more quickly by capturing index criteria from a document's contents. Texts are read by RECOGNITION with the help of OCR (Optical Character Recognition) or from pre-defined zones via barcodes. A broad palette of configuration possibilities make RECOGNITION a multi faceted tool.

Applications

To store incoming invoices, start by sticking barcode labels on a batch of invoices and scan them into DocuWare. Next, the documents are placed into a DocuWare basket, where the barcode numbers are automatically read with the help of RECOGNITION. This module extracts the barcodes as index words for the overall storage process. Now you can record or book the invoices in your accounting program, by entering the company name, invoice number, amount, etc. This data can also be added automatically to the invoice number to enhance the indexing of an invoice in DocuWare — it's all automatic with the add-on module DocuWare AUTOINDEX. If you have a large quantity of paper documents that are set up in a similar way, e.g. forms, you can use OCR to read through text elements that are placed in the same spot, and then incorporate the information to index automatically in DocuWare. If you need to file away work orders and reports generated by service technicians, RECOGNITION makes it

easy. A form is created, for example, in an order processing program in your customer service department. After a request comes in, the form is partially completed with the customer name, order date etc. The order number is added as either a barcode or in a typed format. The area for the service report remains empty. By printing the service report form it is automatically filed away by TIFFMAKER, which uses the customer name, order date, and order number as index words.

By using DocuWare, your customer service department can have all of the orders stored in one file cabinet and can maintain an overview of the status of all the orders.

If a service technician drives to the customer, he takes along a print-out of the partially completed form, fills out any empty fields and has the client sign off on everything. Back at the office, the document is scanned, the barcode is read per RECOGNITION and the document is indexed and securely stored using the index criteria of the formally stored partially complete version of the form.



Highlights

- Text and Barcode Recognition
- Automatic Indexing
- Document Classification (Pre-Sorting)
- Easily Configured with a Handy Editor
- Saves Time — Streamlines Work Processes



Functionality

Incoming paper documents such as invoices, proof-of-delivery or shipping slips and records are captured with a scanner and brought into DocuWare. They are first placed in a DocuWare basket, the reading process then occurs in one of the following ways:

- RECOGNITION is called up from the main DocuWare window and then reads the marked documents from a basket.
- RECOGNITION itself is launched. The reading process can then be run on multiple baskets.
- Together with DocuWare ACTIVE IMPORT, RECOGNITION extracts the index words from the files that are being imported.

With the help of text or barcode recognition, information such as the barcode number, the sender or the date are automatically extracted from the document and then used as index words for storing the document in DocuWare. These are written into the header of the document file. In that way, the document is pre-indexed. Now its ready

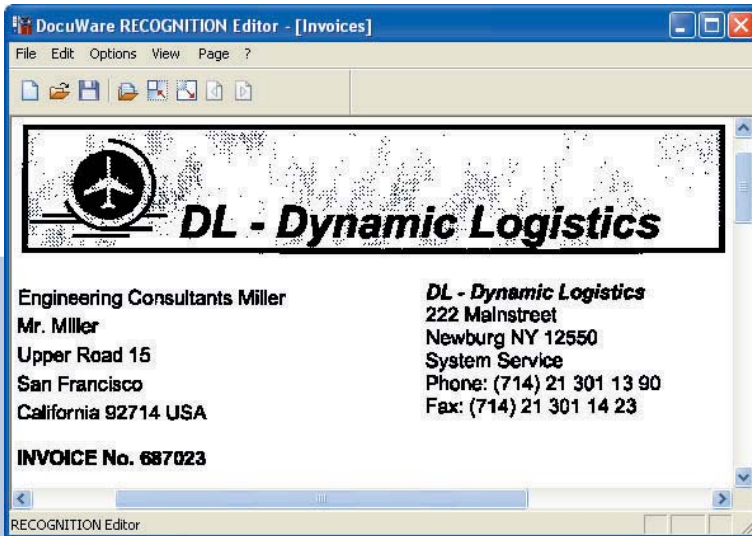
to be moved from the basket into the appropriate file cabinet. At the same time, the index words from the header are automatically copied into the appropriate database fields of the file cabinet.

In combination with ACTIVE IMPORT, you can use RECOGNITION to prepare documents to be imported. Depending on the results of the reading, the documents can be placed in various baskets or into file cabinets, so that document classification (pre-sorting) takes place automatically. DocuWare RECOGNITION is made up of the RECOGNITION program that is responsible for reading/extracting the information, plus the RECOGNITION Editor. Before you work with RECOGNITION for the first time, you need to set up a configuration in the Editor.

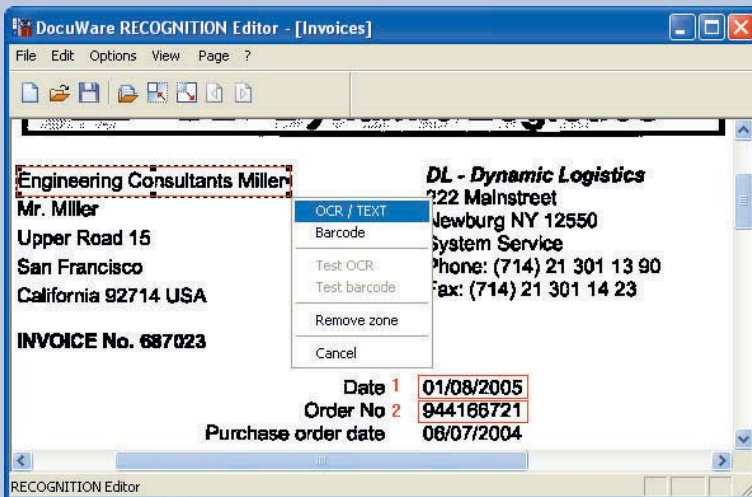
Setting Up a RECOGNITION Configuration

With a new RECOGNITION configuration, you begin by selecting the file cabinet where documents will be stored after the index words have been extracted. You also provide a name for the configuration and select a sample document. This is displayed in the main window of the RECOGNITION Editor.

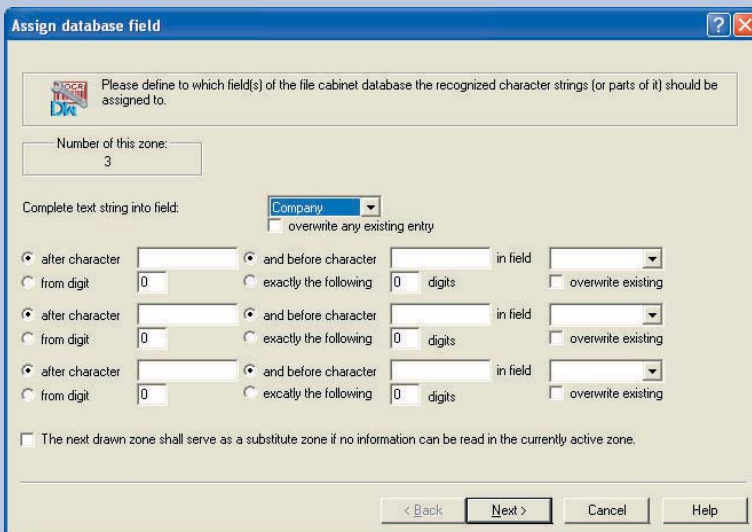
For the configuration, you determine if and how documents should be separated, which is important when scanning stacks of documents. You can also decide if the entire text of a document should be read for full-text indexing. If your scanned documents are not good quality, there are also several options for improving the image quality. For example, if grey areas or noise/lines need to be eliminated in some cases. It is possible to store an enhanced document, replacing the original. For full-text indexing, you enter an OCR setting that is adjusted for your documents to make sure you get the most accurate read. You can also opt to have a log file created to catalog the reading process. And you can also choose fixed index entries for up to five database fields. These can be selected whether or not they are present in the sample document.



Sample document for a RECOGNITION configuration



Defining zones to be read



Assigning index fields

Defining the Areas to be Read

By pulling a rectangle across a portion of the sample document, you determine which areas will be read later.

For each of these areas, you determine whether text (including numbers) or a barcode will be read. You also pick the database field where the text string read from the document will be placed when you store it.

To optimize reading, each area can be fine-tuned through image enhancement. Special OCR or barcode settings can also be adjusted for each area. With all of these choices, you can adapt the RECOGNITION process completely to your needs. The completed RECOGNITION configuration is now available for all DocuWare users in your network.

Functions

General Functions

- Automatic text and barcode recognition for extracting index words
- The index information is read from pre-defined image areas and transferred to assigned database fields

RECOGNITION Editor

- Definition of up to 24 barcode and/or text field entries per file cabinet
- Entry into memo field of up to 64 KB
- Alternative positions definable per field (substitute rectangles)
- Up to 5 fixed or constant field entries
- Separation of documents scanned in a stack with automatic recognition of the first or last page of a document or with the help of separator pages in a batch
- 100% barcode recognition with sufficient quality

Barcode Types

- The width of the narrowest line must be a min. of 0.25 mm
- Code 39: alphanumerical, relatively wide, variable length

- CODABAR: numerical, error recognition characters, variable length
- UPC/EAN: internationally used for commodities, numerical, fixed length, error recognition and check sums, relatively narrow
- UPC 6: used in USA, 6 characters long
- UPC 12: used in USA, 12 characters long
- EAN 8: used internationally, 8 characters long
- EAN 13: used internationally, 13 characters long
- 2 out of 5: variable numerical length, generally compact, easy to recognize: Datalogic, Interleaved, 5 Line and Matrix

OCR-Settings

General:

- Use of standard OCR settings which usually generate good reading results
- Customize special OCR settings
- Various OCR settings can be set globally for an entire configuration or individually for each field
- OCR settings can be stored in multiple configurations

Languages:

- OCR for special international characters (Western and Eastern European, Baltic, Cyrillic, Turkish, Greek and Maltese character sets)
- Select the language(s) which will be used in the documents, to limit the character set (34 languages available)
- Select level of recognition accuracy and influence the speed of the reading process

Character Filters:

- Reducing the number of characters to reduce time for recognition process

Image Enhancement:

- Automatic image enhancement: documents with a poor quality can be altered to improve reading results
- With image enhancement for an entire document, the new document can be stored in lieu of the original
- Straighten or rotate images

For more information
please visit our website at
www.docuware.com