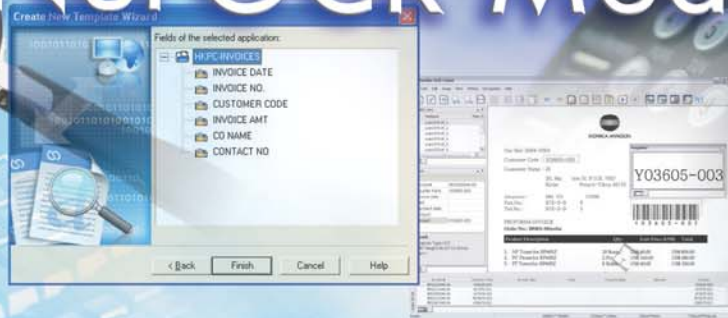


# DocNet OCR Module



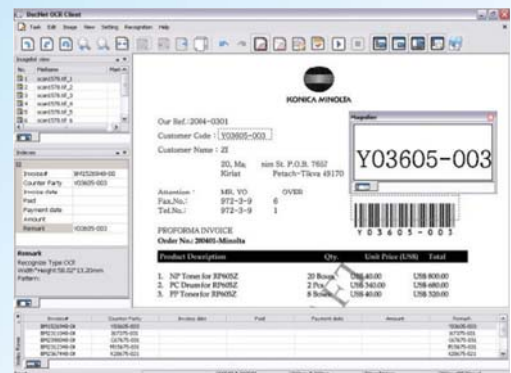
How do you give indexes to tedious of documents in DocNet? Keep typing the indexes for each document one by one? DocNet OCR Client Module gives you a hand on such scenario.

With the installation of DocNet OCR Client Module and scanners, PC will turn to a scanning station for DocNet immediately. Users are able to create different templates for different documents, designate recognition regions on the templates. Once the templates are formed, documents can be scanned via the scanner to a designated hot folder, and indexes can be extracted automatically with a single click. Finally, images and indexes will be uploaded to the DocNet server and ready for retrieval.

## What is DocNet OCR Client Module?

The DocNet OCR Client Module comes with 2 recognition engines: the Optical Character Recognition (OCR) and Barcode Recognition (Code 39), which handle text and barcode recognition respectively.

In general, the auto indexing involves 2 major operations. Firstly, users are required to create a template for defining the areas in a certain document for recognition. Secondly, once the template is set up, batches of documents can be proceeded by a single recognition task using the predefined template. Indexes and images will then be uploaded to the DocNet Server after the recognition process. With the DocNet OCR Client Module, documents such as invoices, purchase order and reports can be recognized efficiently.



■ Create a new template for image data file recognition

## Templates

Template is a key component for the DocNet OCR Client Module, which provides a platform to define different areas in a document for index extraction. Different image formats are supported such as TIFF, JPEG and BMP to be imported as a template. Connection to DocNet Server provides a list of index fields of any specific application, and field attributes are displayed for reference. Users can simply drag different areas for both OCR and barcode recognition.

**DocNet**  
your Document Imaging Solution



**KONICA MINOLTA**

The essentials of imaging

Enquiry Hotline: 2565 2111

<http://konicaminolta.com.hk>

## High Accuracy of Recognition

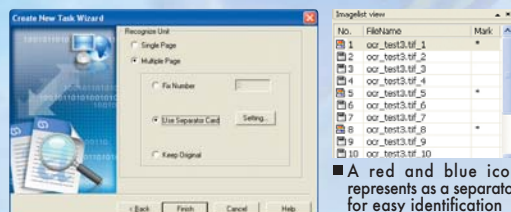
Users are able to apply patterns to each index fields for the optimal accuracy. For instance, invoice number like BM-14704, users are able to create a pattern `**#####` where `*` means uppercase characters and `#` means numbers. Suspected strings can be judged by this mechanism, and the error rate can be minimized. Such setting can speed up the recognition process and increase the accuracy of recognition.

## Batch Scanning Task

Scanning documents is no longer an exhausting job. A stack of documents can be scanned rapidly by means of the FTP or SMB provided by our MFPs. The DocNet OCR Client Module is able to monitor a document folder for a certain task. Documents arrives the document folder and the recognition process is ready to launch. Users are able to start the recognition process and check the results simultaneously. Results will be displayed for validation before the upload-to-DocNet process.

## Multiple Page Documents Handling

How to batch scan a stack of papers that contains several sets of document? The DocNet OCR Client Module provides separating function that divides a stack of papers into several sets of documents. Fixed pages, separator pages can be treated as an identifier for dividing pages. Pages with special characteristics such as a company name or different page size can be used as a separator page. It can be defined in the DocNet OCR Client Module easily, and efficiency of scanning documents can be increased immediately.



■ By selecting the recognize unit, you can separate the image file into different sets of documents.

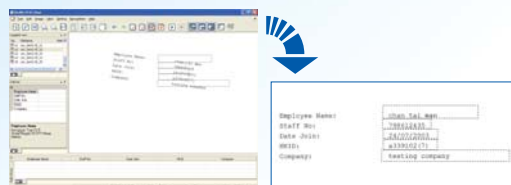
■ A red and blue icon represents as a separator for easy identification

## Multi DocNet OCR Client Module Environment

In case there is a busy application that stores lot of documents scanned by different MFPs daily, it is possible to have several DocNet OCR Client Modules for serving this need. A single share folder can be setup among the DocNet OCR Client Modules to collect different image files scanned by different MFPs. Individual DocNet OCR Client Module are able to get an image file from the share folder and name it with a file extension of `*.processing`. This is a sign to the other DocNet OCR Client Modules that this image file is being handled by others. Thanks to this mechanism and multitasking of image processing can be done effectively.

## Image Editing

Image quality sometimes affected by misfeeding of scanner or the condition of paper. In order to eliminate the error caused by the poor image quality, the DocNet OCR Client Module provides deskewing for compensation of ruin images. In addition, the noise and erode settings are provided for image enhancement. These procedures can be done before the recognition process to improve the recognition accuracy, which is a convenient way to avoid the re-scanning of images.



■ After deskew, images can be corrected without re-scanning

With its versatile features, DocNet OCR Client Module is the ideal solution for your auto filing and indexing!

### System Requirements

#### Hardware

- Pentium 4 2.0GHz or higher
- 512 MB RAM
- 40GB HDD
- 8MB VGA Card
- 10/100 Mbps TCP/IP Network
- 1 x USB slot

#### Software

- Windows 2000, Windows XP Professional

**DocNet**  
your Document Imaging Solution

  
**KONICA MINOLTA**