



# Taskmaster Flex

## Quick Start Guide

## **Copyright Notice**

©Copyright 1995-2009 Datacap Inc.

### **All Rights Reserved.**

No part of this publication may be copied, photocopied, reproduced, transmitted, transcribed, or reduced to any electronic medium or machine-readable form without the prior written consent of Datacap Inc.

U.S. GOVERNMENT RESTRICTED RIGHTS. The Software and Documentation are provided with RESTRICTED RIGHTS. Use, duplication, or disclosure by the Government are subject to restrictions set forth in subparagraph (c)(1)(ii) of the Rights in Technical Data and Computer Software clause at DFARS 252.227-7013 or subparagraphs (c)(1) and (2) of the Commercial Computer Software Restricted Rights at 48 CFR 52.227-19, as applicable. Manufacturer is Datacap Inc., 660 White Plains Road, Tarrytown, NY 10591.

Printed in the United States of America

## **Trademarks**

**Microsoft Access® is a registered trademark of Microsoft Corporation.**

**All trade names and marks used herein are the property of their respective owners.**

### **Datacap Inc.**

660 White Plains Road  
Tarrytown, New York 10591  
Tel: 914-366-0100  
Fax: 914-366-0500

Email (Information): [info@datacap.com](mailto:info@datacap.com)

Email (Support): [support@datacap.com](mailto:support@datacap.com)

Internet: <http://www.datacap.com>

## Table of Contents

<b>Taskmaster Flex</b> .....	<b>4</b>
<i>Introduction to Taskmaster Flex</i> .....	4
<i>Introduction to Taskmaster Flex Quick Start Guide</i> .....	5
<i>Before You Begin</i> .....	5
Taskmaster Flex and Taskmaster FlexID Prerequisites .....	5
Make Backup Copy of c:\Datacap\Flex Folder .....	5
Define Data Source Name (DSN) .....	5
Start Taskmaster Server Service.....	7
<i>Using Taskmaster Flex</i> .....	8
Tasks Performed by Demo Application .....	8
Run VScan Task.....	9
Run Rulerunner Task.....	10
Run Verify Task .....	10
Run Export Task.....	14
Re-run Demo Job to Show Auto-Classification .....	14
<i>Using Flex Configurator</i> .....	15
Index Fields Tab.....	15
Document Class Tab .....	18
<i>Set Up Taskmaster Flex to Process Your Document Images</i> .....	18
Delete Sample Input Images.....	18
Copy Your Images to Input Folder.....	19
Define New Document Classes and Index Fields .....	19
Flex Configurator Field Definitions .....	20
<b>Taskmaster FlexID</b> .....	<b>24</b>
<i>Introduction to Taskmaster FlexID</i> .....	24
<i>Using Taskmaster FlexID</i> .....	24
Multi-Page TIFFs .....	24
Setup Document Classes and Index Fields in Flex Configurator .....	25
Tasks Performed by Demo-FlexID Job.....	25
Run VScan Task.....	26
Run FlexID Task .....	26
Run Rulerunner Task.....	27
Run Verify Task .....	27
Run Export Task.....	28
Re-run Demo-FlexID Job to Show Auto-Classification.....	29
<i>Deleting Taskmaster Fingerprints</i> .....	29
Datacap Studio .....	29
The Zones Tab.....	31
Deleting Fingerprints.....	32
Viewing Document Zones.....	32
<b>Default Index Fields and Field Definitions</b> .....	<b>34</b>

# Taskmaster Flex

## *Introduction to Taskmaster Flex*

Taskmaster Flex provides easy document identification and data extraction. It raises data capture to a new level of simplicity. Build an application in minutes by providing Taskmaster Flex with the Document Class (type of document) and index field names. Taskmaster Flex searches for the fields, and automatically retrieves the data. If the field cannot be automatically found, Taskmaster Click N Key enables you to populate data fields by pointing to the data and clicking. That's it! No programming, no template building, and no workflows to configure.

Taskmaster Flex is smart! It remembers the Document Class, fields, and field locations. The next time the same document is scanned, Taskmaster Flex automatically identifies the document, extracts the required data fields, presents the data and the document image for verification, and exports the data in XML format for input into your document management or imaging system. Taskmaster Flex is truly "Easy OCR." You can provide optional data cleaning and validation information to complete the application.

Taskmaster Flex features include:

- **Quick and easy setup.** Preconfigured workflow for indexing eliminates application development and testing time, giving you the fastest solution for developing data capture applications.
- **Proven DNA technology.** Taskmaster Flex is built on Datacap's advanced Dynamic Natural Analysis (DNA) techniques such as Intellocate, which automatically searches for and retrieves field data using labels and zones, and on Click N Key which lets you quickly identify field data by pointing and clicking. DNA eliminates most document setup and enables single-pass verification for fast and accurate data capture.
- **One-file setup.** A single settings file defines all database connections and lookups, behaviors, even the graphics displayed on the screen. This greatly simplifies configuration and deployment, and enables you to maintain a test and production system using exactly the same code base by using different settings files.
- **Flex Configurator.** This simple GUI tool enables an administrator to quickly and easily define Document Class, fields, and optional data validations. No programming and no Taskmaster experience is required, making setting up a Taskmaster Flex application something virtually anyone can do without training.
- **XML output.** The default tagged XML output contains all necessary extracted data and the location of each image file, making integration with back-end systems a snap.
- **Self-learning.** Given a list of indexes, Taskmaster Flex learns not only which indexes are appropriate for each document, but also where those fields are located on that document. It performs OCR recognition and populates the results automatically.
- **Build-in validations.** With absolutely no programming, Taskmaster Flex provides validation of data types, data lengths and formats, and filtering of data input.
- **Extensible.** The entire application is built using the standard Taskmaster architecture. Changes and additions are implemented using familiar Taskmaster rules, no special programming needed.

Taskmaster Flex also includes Taskmaster FlexID. Taskmaster FlexID provides simple page identification without the use of separator pages. It displays thumbnails of document pages and enables you to specify whether a page is the beginning of a new document, the

continuation of the previous document, or an attachment to the previous document. To learn about the capabilities of FlexID and how to use it, see the section titled **Taskmaster FlexID**.

### ***Introduction to Taskmaster Flex Quick Start Guide***

The purpose of this Taskmaster Flex Quick Start Guide is to introduce you to Taskmaster Flex and Taskmaster FlexID, to walk you through the major features and the capabilities of both, and to get you started using the out-of-the-box sample jobs as quickly as possible.

This Taskmaster Flex Quick Start Guide provides the information you need in order to run Taskmaster Flex Windows-based (not Web-based) client tasks in an existing Taskmaster 7.5 SP1 or later environment where all components are installed on a single machine (referred to in these instructions as the Workstation).

Included with Taskmaster Flex are sample document images and Document Classes (types of documents) defined. These samples are included so you can use Taskmaster Flex out-of-the-box without additional setup steps required.

These instructions assume you are running Taskmaster Flex with all components installed on a single machine *for demonstration purposes only*. For information about installing, configuring, and running Taskmaster applications in a client/server, Windows- or Web-based environment (recommended for development, test, and production environments), see the Taskmaster 7.5 SP1 Installation and Configuration Guide found in the root folder on the installation media or in the c:\Datacap\Docs folder.

### ***Before You Begin***

Ensure you have met the prerequisites listed here before you begin using Taskmaster Flex and Taskmaster FlexID.

#### **Taskmaster Flex and Taskmaster FlexID Prerequisites**

Please ensure you have a working Taskmaster 7.5 SP1 environment on a single machine. At a minimum, ensure at least one of the sample Taskmaster applications runs correctly, as this makes troubleshooting Taskmaster Flex easier, and narrows down potential problems that may occur.

By default, Taskmaster Flex and Taskmaster FlexID are installed when you perform a complete installation of Taskmaster 7.5 SP1.

Ensure you have administrator access to the machine on which Taskmaster software is installed.

#### **Make Backup Copy of c:\Datacap\Flex Folder**

Use Windows Explorer to navigate to and make a copy of the c:\Datacap\Flex folder, ensuring the copy is stored outside of the Datacap folder.

#### **Define Data Source Name (DSN)**

There are six Microsoft Access databases in the c:\Datacap\Flex \Process folder. You must create a DSN on your machine that Taskmaster Flex can use to connect to the Microsoft Access database (.mdb) that contains the Flex fingerprints. The DSN name used in the connection string in the Flex application is configured as “Flex application name” concatenated with “Fingerprint.” For Taskmaster Flex, as it is delivered out-of-the-box, you must create a DSN named “FlexFingerprint.”

## Quick Start Guide

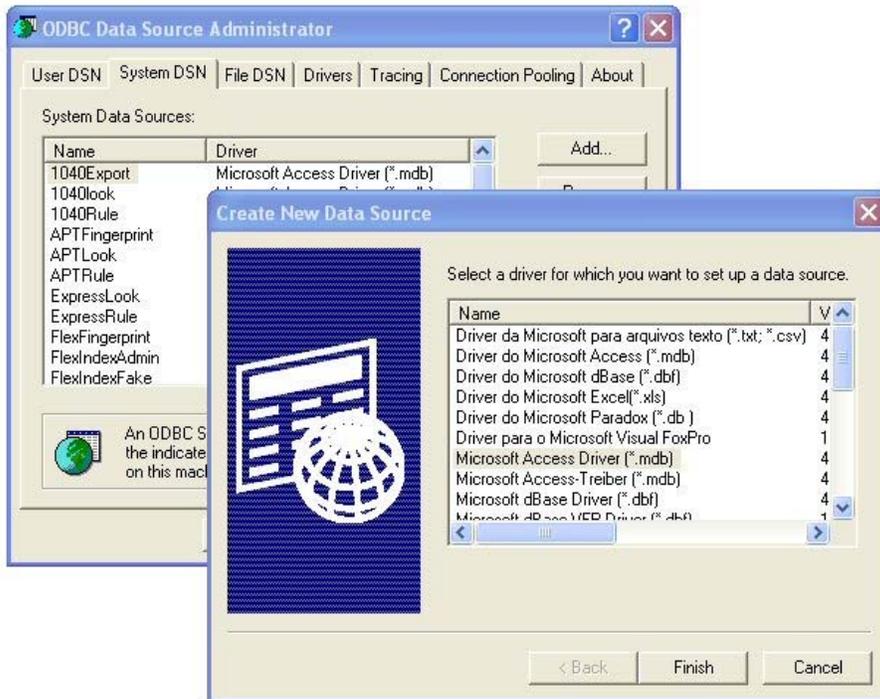
---

Here are the steps to take to create the DSN:

From your Windows **Start** menu, select **Control Panel**, select **Administrative Tools**, and click **Data Sources (ODBC)**.

Select the **System DSN** tab and click **Add**.

Select **Microsoft Access Driver (\*.mdb)**



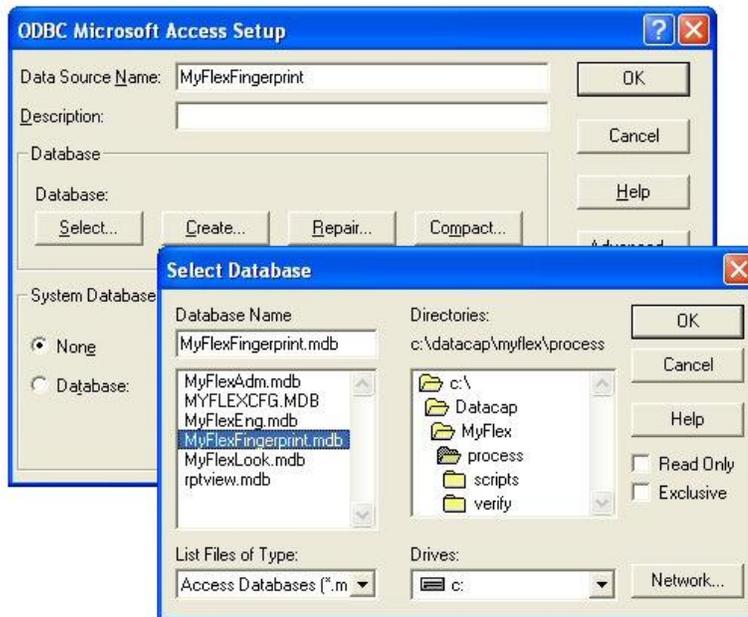
Click **Finish**.

Enter your **Data Source Name**: FlexFingerprint

Click **Select** and locate the FlexFingerprint.mdb file in the Process folder in your Flex application folder:

c:\Datacap\Flex\process

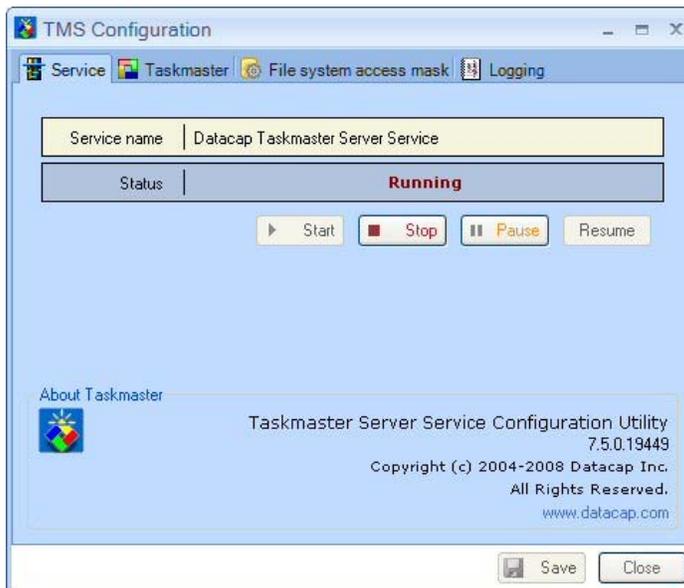
Highlight the FlexFingerprint.mdb file and click **OK**.



Click **OK** to close the ODBC window.

### Start Taskmaster Server Service

From your Windows **Start** menu, select **All Programs**, select **Datacap**, select **Taskmaster Server**, and click **Taskmaster Server**. The TMS Configuration window opens:



The status should be *Running*. If not, click **Start**. A Starting Service message is displayed. The Status changes to *Running* when the service is started. Close the TMS Configuration window.

### Using Taskmaster Flex

Taskmaster Flex is shipped with four sample document images and with three Document Classes (types of documents) defined.

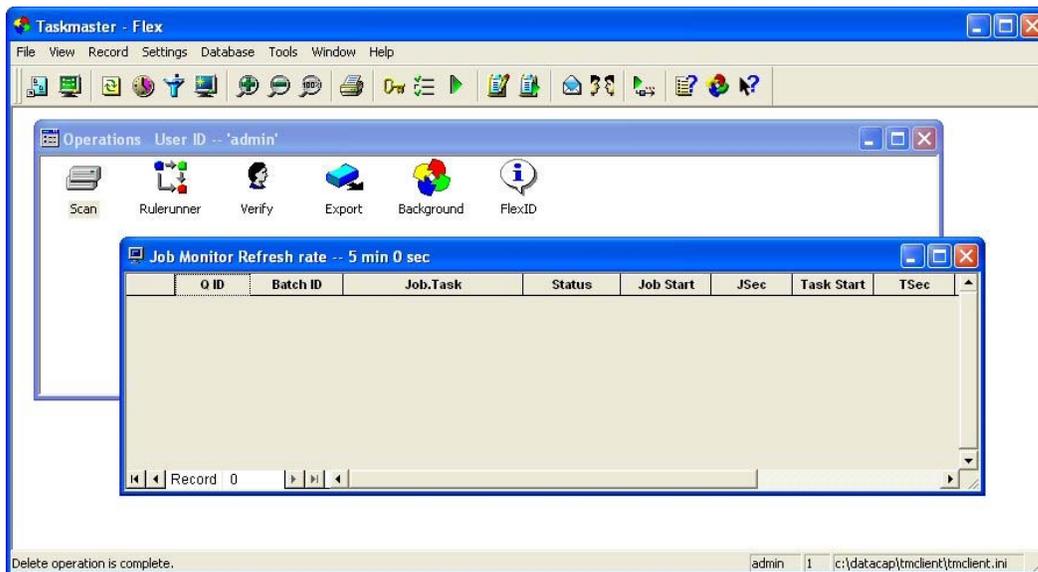
The Document Classes are for an invoice, tax return, and email message. The Document Classes and the fields for the documents are stored in a Microsoft Access database which is part of the sample application. Document classes and fields are defined using the easy-to-use Flex Configurator tool, which is described later in this document.

The sample images are in c:\Datacap\Flex\images\Input

To run the Taskmaster Flex Client, from your Windows **Start** menu, select **All Programs**, select **Datacap**, select **Applications**, select **Flex** and click **Flex Client**, or click the application shortcut, **Flex Client** in c:\Datacap\Flex.

Enter a server address of **127.0.0.1** and **admin** for the User ID and Password. The Taskmaster Flex application starts, and displays the Operations window.

To display the Job Monitor window, on the **View** menu, select *Job Monitor*.



### Tasks Performed by Demo Application

Taskmaster Flex has several different sample “jobs” defined. A job is a collection of tasks to perform and rules to control document processing.

The *Demo Job* is made up of four tasks:

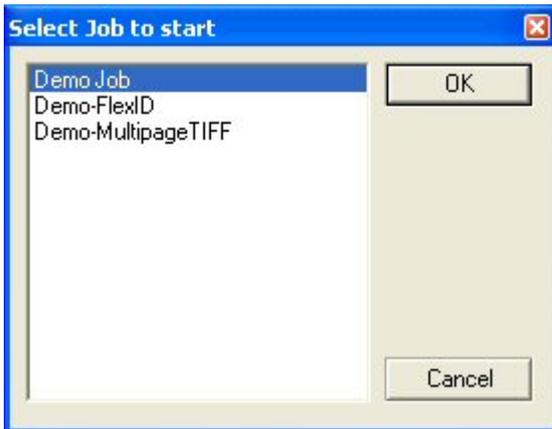
- **VScan** – This virtual scan task brings the sample document images into Taskmaster Flex.
- **Rulerunner** – This task performs many image processing functions including identification of pages, image cleaning (e.g. despeckling), image rotation (if necessary), full page character recognition, and more. It also attempts to identify the document by comparing it to a library of known document types.
- **Verify** – This task displays the documents in the Verify panel. If the document is not known to Taskmaster Flex, you identify the document type by selecting it from a dropdown list, and then identifying where the data fields are located. If the document is

known to Taskmaster Flex, the Verify panel displays the fields and data so you can ensure they are correct.

- **Export** – This task saves information about the documents and exports field data to XML files for input into a back-end application such as a document imaging or ERP system.

### Run VScan Task

Double-click the **Scan** icon, select **Demo Job** and click **OK**.



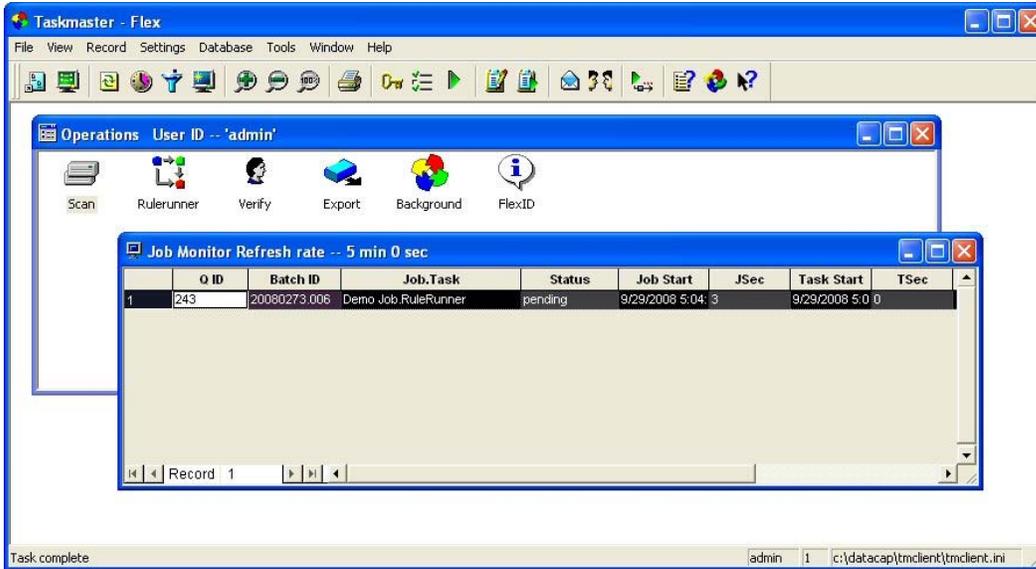
A Rulerunner status window briefly appears as the images are processed. When processing is complete, a message box is displayed indicating the task has finished. Click **OK**. Another message box asks “Do you want to continue processing?” Click **No**.

Taskmaster Flex has picked up the images from the folder `c:\Datacap\Flex\images\Input` and created a Taskmaster “batch,” a collection of document images to be processed. It assigned a batch ID and created a folder with the name of the batch ID in `c:\Datacap\Flex\batches`. This batches folder holds all of the document image processing results for this batch of documents.

If the Job Monitor window is not already open, open it by clicking the **View** menu and selecting *Job Monitor*. The contents of the Job Monitor window are refreshed at a pre-defined interval which is five minutes by default. You can refresh the Job Monitor window immediately by clicking it and pressing F5.

### Run Rulerunner Task

After running the VScan task, run the Rulerunner task by double-clicking the **Rulerunner** icon in the Operations window, or double-clicking the batch number (e.g. "1") of the batch you want to run in the Job Monitor window.

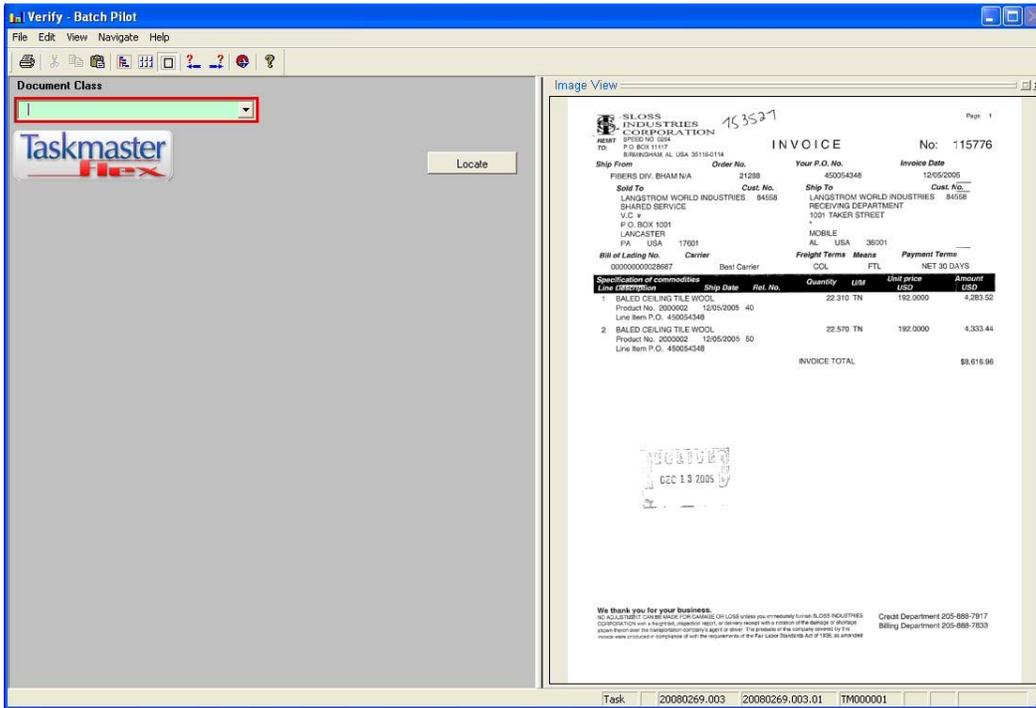


The Rulerunner status window is displayed as the document images are processed. This may take a few minutes, as Rulerunner is running a number of processing-intensive modules, including PageID, ImageFix, Recognition, lookups, Locate, and others.

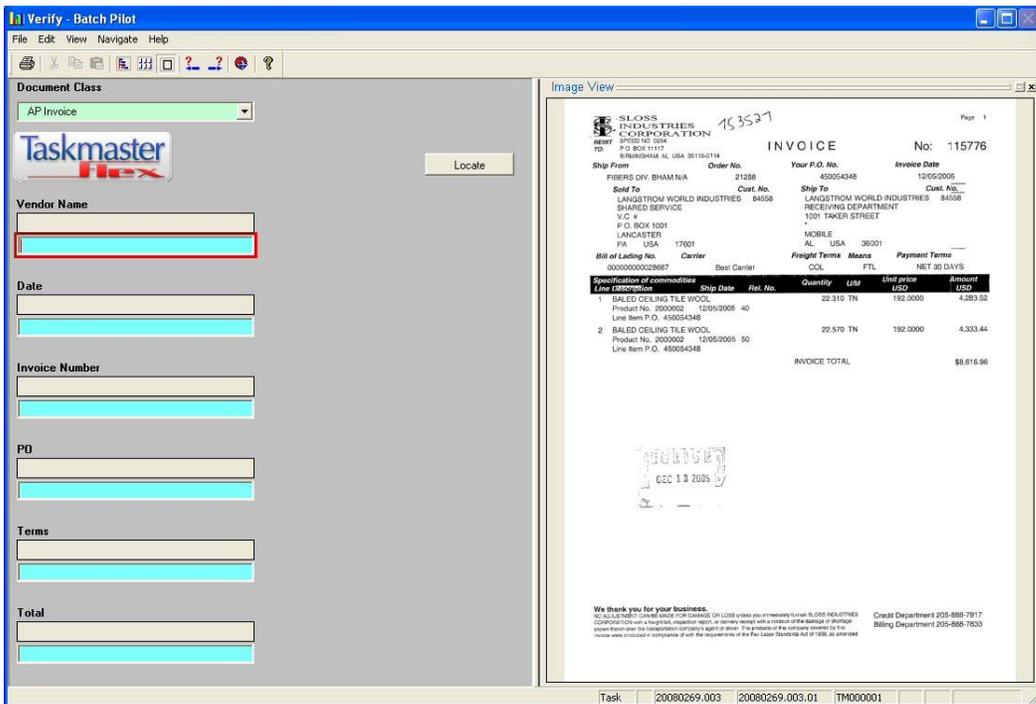
### Run Verify Task

When the Rulerunner task completes, refresh the Job Monitor window, and run the Verify task by double-clicking the **Verify** icon in the Operations window, or double-clicking the batch number of the batch you want to run in the Job Monitor window.

The Verify task displays the document image and a **Document Class** field from which you select the type of document.



For the first document in the batch, select the **AP Invoice Document Class**. *AP Invoice* is one of the three default Document Classes provided with Taskmaster Flex. Once you select the **AP Invoice Document Class**, six fields are displayed.



Click the **Locate** button to direct Taskmaster Flex to search for field values on the document image. Using built-in location rules, Flex finds Date, PO and Total. You must type in the

## Quick Start Guide

Vendor Name as the vendor name in the logo on the image cannot be recognized as character text by the recognition engine.

### Note on Locate Button

The **Locate** button finds values for fields that have been pre-defined in Taskmaster Flex with a Locate Method of *Keyword File* or *Regular Expression*. The Locate Method for a field is defined on the **Index Fields** tab in the Flex Configurator. Click the **Locate** button to find field values the first time a new document type is viewed in the Flex Verify panel.

Now use Taskmaster Click N Key to identify the other fields on the image. To do this, first click inside an empty data entry field and then point at the field value on the document image and click it, or click and drag the mouse to draw a rectangular box around the field on the image. Do this for each of the fields that do not yet contain data.

The completed Verify panel should look like:

The screenshot shows the Taskmaster Flex Verify Batch Pilot interface. The left pane contains a 'Document Class' dropdown set to 'AP Invoice' and a 'Locate' button. Below this are several data entry fields, many of which are populated with values from the invoice image:

- Vendor Name:
- Date:
- Invoice Number:
- PI:
- Terms:
- Total:

The right pane, titled 'Image View', displays an invoice from LANGSTON WORLD INDUSTRIES. The invoice includes a table of line items:

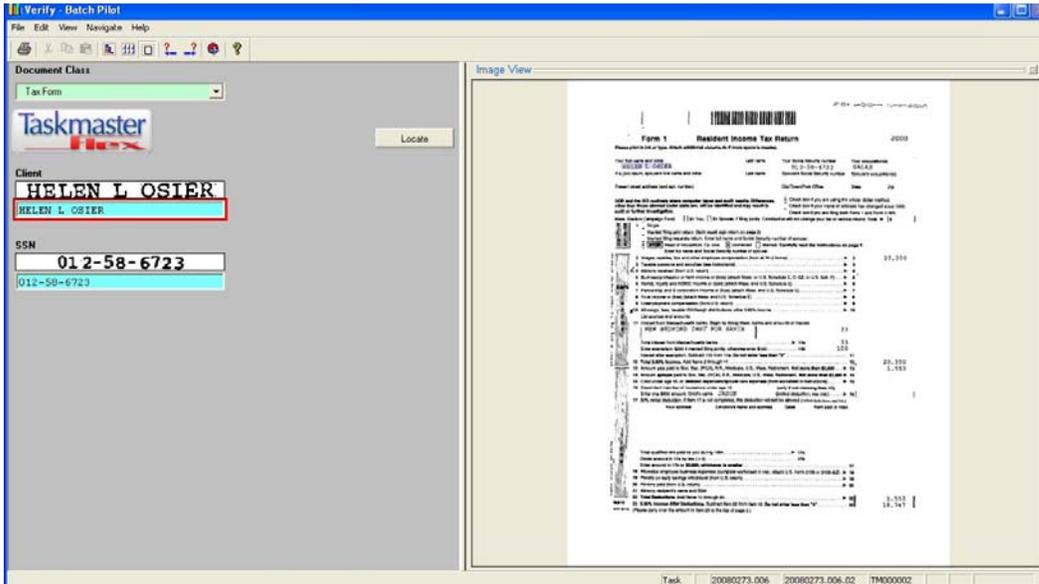
Line Description	Ship Date	Ref. No.	Quantity	UNIT	List Price	Amount
1 BILED CEILING TILE WOOL			22.910	TN	192.0000	4,389.52
2 BILED CEILING TILE WOOL			22.970	TN	192.0000	4,409.44
INVOICE TOTAL						\$8,798.96

When you have captured all of the field data, click the **Next**  arrow button to go to the next document in the batch, which is a U.S 1040EZ tax return.

Select the **Tax Form Document Class**.

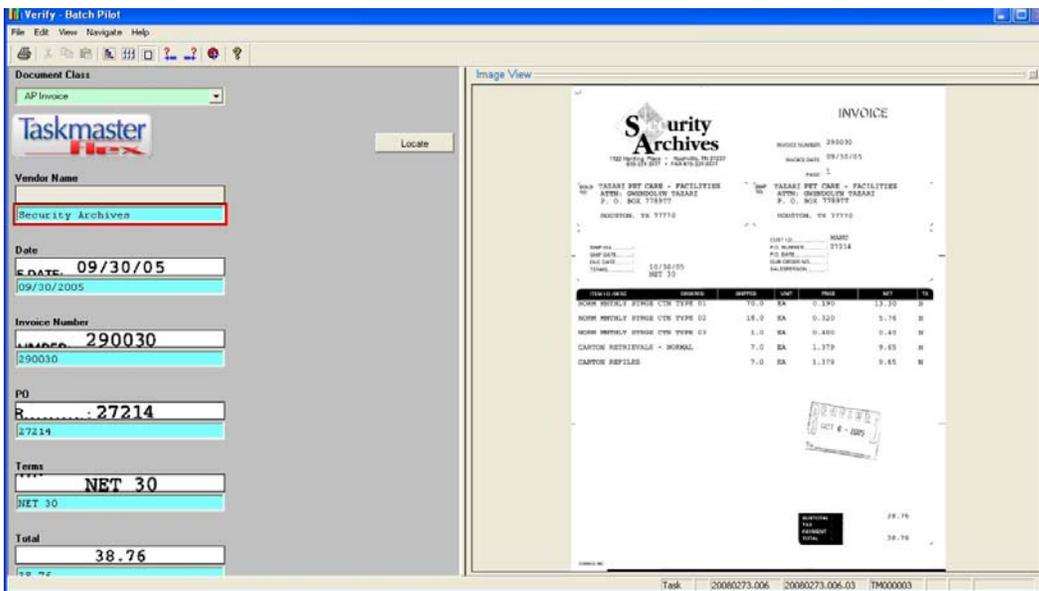
Click the **Locate** button, and Taskmaster Flex finds the **Social Security Number**.

Click in the **Client** field, then click and drag the mouse to draw a rectangle around the name “Helen L Osier” on the image. The Verify panel should now look like this:



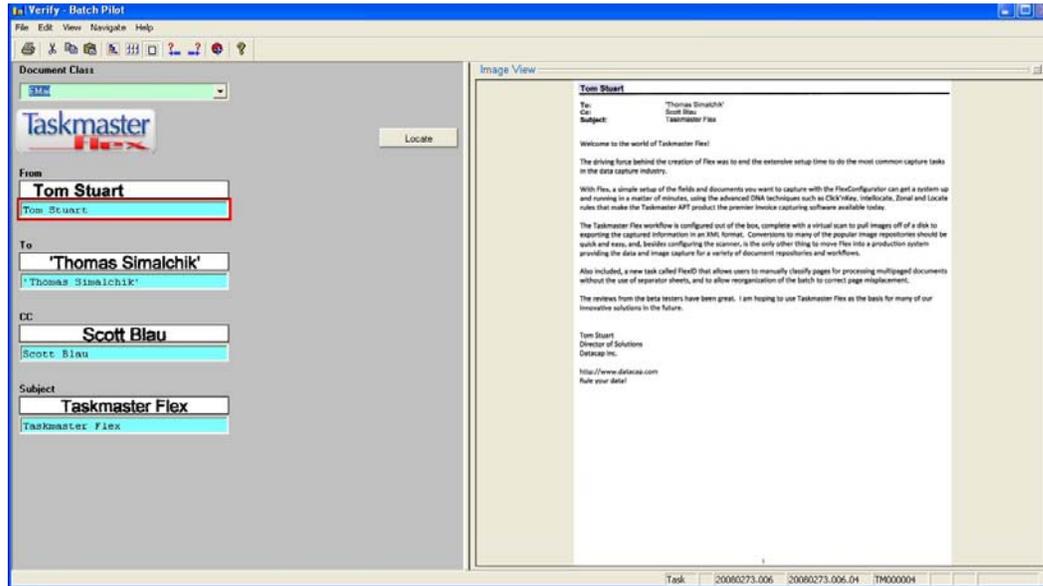
When done, click the **Next**  arrow button to go to the next document image, which is another invoice. Select **AP Invoice** for the **Document Class**, click **Locate**, and then identify the remaining fields. Type in the name **security Archives** for the **Vendor Name** field. If you use Click N Key to identify the **PO** number, notice that the preceding period may be picked up: **.’27214** – delete the period. Taskmaster Flex remembers that you performed an edit on the field so when you run Flex again with these same images, the PO number is located and displayed without the period.

The Verify panel should now look like this:



Click the **Next**  arrow button to go to the last document image which is an email.

Select *Email* for the **Document Class**, click **Locate**, and then identify the blank **From** field. The Verify panel should now look like this:



Click the **Next**  arrow button and a message box is displayed indicating “No more problems have been found. Click OK to finish the batch.” Click **OK** and another message box is displayed asking if you want to continue processing. Click **No**. The Verify panel closes.

### Run Export Task

When the Verify task completes, refresh the Job Monitor window, and run the Export task by double-clicking the **Export** icon in the Operations window, or double-clicking the batch number of the batch you want to run in the Job Monitor window.

Export performs two important functions:

- Export saves information about the documents and their fields so that the next time Taskmaster Flex “sees” the same document, it can identify the document type and extract the field data. (Flex accomplishes this by creating a Taskmaster “fingerprint.”)
- Export outputs the field data to XML files. The XML files and the image files are now ready to be input into a back-end application such as a document imaging system.

When Export completes, refresh the Job Monitor window. The status for this batch of documents is now *Job Done*.

You have trained Taskmaster Flex to identify four documents and extract data from the documents without any programming. The next time Flex sees one of these documents, it knows the document type and where to find the field data.

### Re-run Demo Job to Show Auto-Classification

To verify that Taskmaster Flex now understands these four documents, you can re-run the Demo Job. Double-click the **Scan** icon, select **Demo Job** and click **OK**. When the VScan task is complete, refresh the Job Monitor window, and double-click the job number for this new batch to run the Rulerunner task. Rulerunner compares these documents against the library of known documents and finds matches. Flex now knows the document type (invoice, email, or tax return) and where to find the data for the required fields.

Refresh the Job Monitor window, and double-click the Job number to run Verify. Notice that Taskmaster Flex now displays the **Document Class** (type) and the fields. The data for the fields is displayed, with the exception of the vendor name for the invoices, which must be typed. (In an actual production application the vendor name and other vendor information would have been extracted from a vendor database.)

Go through and view each document by pressing the **Next**  arrow button. Taskmaster Flex has “learned” these documents and can now automatically determine the **Document Class** and extract the required field data. Exit when complete. You do not need to run the Export task.

### Using Flex Configurator

The Flex Configurator is a GUI tool that enables an administrator to quickly and easily define and update Document Classes, fields, and optional data validations. No programming or Taskmaster experience is necessary. Setting up new documents with Taskmaster Flex is something virtually anyone can do without training.

To run the Flex Configurator, from your Windows **Start** menu, select **All Programs**, select **Datacap**, select **Applications**, select Flex, and click Flex Configurator.

### Index Fields Tab

The Flex Configurator has two simple tabs; by default, the Index Fields tab is displayed on top.



There are twenty pre-defined index fields that can be used to index a document in the Flex sample application. For a complete list of the index fields delivered with Taskmaster Flex and their definitions, see **Default Index Fields and Field Definitions**.

## Quick Start Guide

These fields are available to be used with any Document Class. In the example above, Address 1 field is selected to show its attributes. Address 1 is alphanumeric without a specified length. A value for the address field is not required. With Find by Zone set to Yes, Flex searches for Address 1 by using the zone, the location on the page, which you identify the first time the document is processed in Flex. Specifically, the zone is the area on the page identified when you use Taskmaster Click N Key or click and drag a box around the data.

Here we see the field definition for Contract No:

The screenshot shows the 'Taskmaster Flex Configurator' window. At the top, it says 'Connection opened successfully' and has a 'Refresh' button. The main area is divided into two tabs: 'Index Fields' and 'Document Class'. Under 'Index Fields', there is a 'Field List' on the left containing various fields like 'Address 1', 'Address 2', 'CC', 'City/State', 'Client', 'Comment', 'Contract No' (which is selected), 'Date', 'Email', 'From', 'Invoice Number', 'Phone', 'PO', 'SSN', 'Subject', 'Terms', 'To', 'Total', 'Vendor Name', and 'Zip'. To the right of the list are configuration fields for the selected field: 'Index Field Name' (Contract No), 'Data Type' (AlphaNumeric), 'Min Length' (7), 'Max Length' (7), 'Character Filter' (empty), 'Picture String' (ANNNNNN), 'Required' (No), 'Find by Zone' (Yes), 'Locate Method' (None), 'Keyword File' (empty), and 'Regular Expression' (empty). There are 'Save' and 'Delete' buttons at the top right of the configuration area.

Contract number is defined as alphanumeric and seven characters in length. The data must match a picture string of ANNNNNN. This picture string requires that the contract number begin with an alpha and be followed by six numerics. Of course, your definition of the format for a contract number can and probably will be different. You can easily change the format to whatever you need. For a definition of all the options for a picture string, as well as all other available validations and attributes, see the section titled [Flex Configurator Field Definitions](#).

Here we see the field definition for an invoice number:

The screenshot shows the 'Taskmaster Flex Configurator' window. At the top, it says 'Connection opened successfully' and has a 'Refresh' button. The main area is divided into two tabs: 'Index Fields' and 'Document Class'. Under 'Index Fields', there is a 'Field List' on the left containing various fields like 'Address 1', 'CC', 'Client', etc., with 'Invoice Number' selected. To the right of the list are several configuration fields for the selected field:

- Index Field Name:** Invoice Number
- Data Type:** AlphaNumeric
- Min Length:** 2
- Max Length:** 20
- Character Filter:** (empty)
- Picture String:** NNNNNN
- Required:** Yes
- Find by Zone:** Yes
- Locate Method:** Keyword File
- Keyword File:** InvNum
- Regular Expression:** (empty)

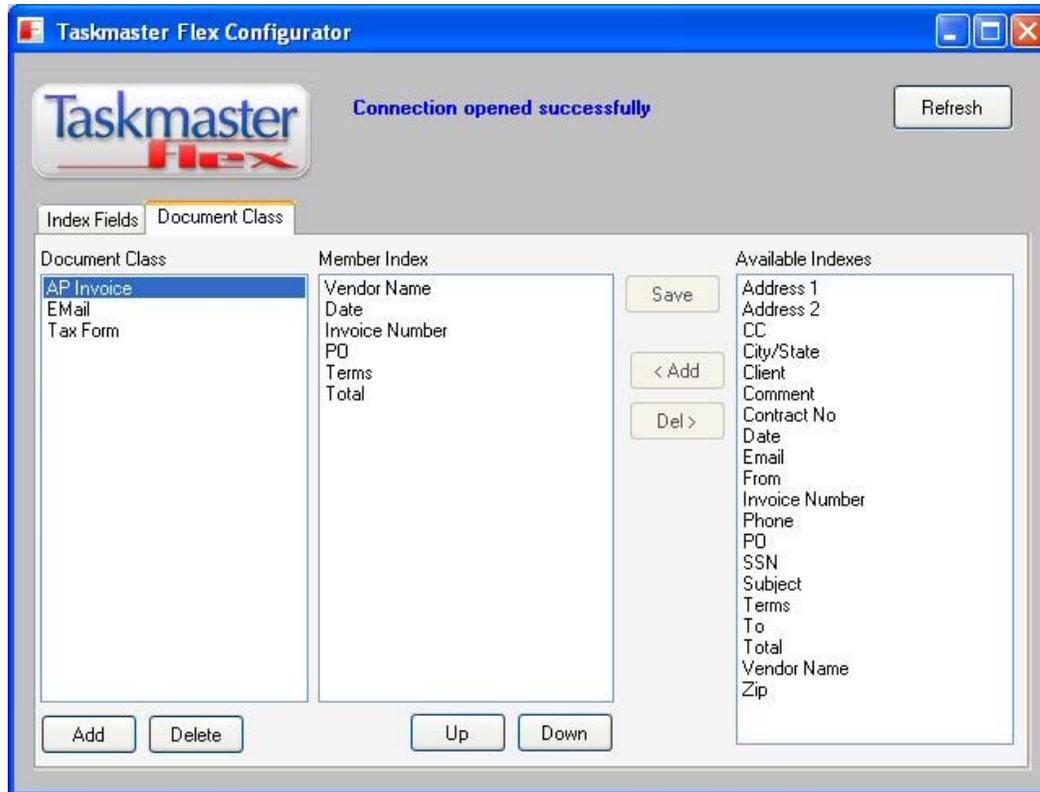
Buttons for 'Save' and 'Delete' are located above the 'Data Type' field.

Invoice number is also defined as alphanumeric with a length of two to 20 characters all numeric. A **Location Method** of *Keyword File* has been specified. A *Keyword File* contains a list of all the possible variations for the ways the label “Invoice Number” can appear on an invoice. For example, it can appear as “INVOICE No:,” “Invoice Number” or “Invoice #:,” among others. Taskmaster Flex uses these variations when it searches for the location of the invoice number. The keywords are stored in a file. Details are provided in the section titled .

To add a new field, type the new field name in the **Index Field Name**, select a **Data Type**, and click **Save**. All other attributes are optional. For a definition of the other attributes, see the section titled **Flex Configurator Field Definitions**.

### Document Class Tab

The Document Class tab is used to add a new Taskmaster Flex and index fields, or change the index fields for an existing Document Class. In this example, the *AP Invoice* **Document Class** has six fields: **Vendor Name, Date, Invoice Number, PO, Terms** and **Total**.



- To add a **Document Class**, click the **Add** button below the list of Document Classes. A dialog box prompts you for the new **Document Class** name.
- To delete a **Document Class**, click the class name and click the **Delete** button below the list of Document Classes.
- To delete an index field for a Document Class, click the field name in the **Member Index** list and click the **Del>** button.
- To change the sequence of the indexes appear on the Verify panel, click the field name in the **Member Index** list and use the **Up** and **Down** buttons to reorder.
- To add an index field to a Document Class, click the index name in the **Available Indexes** list and click the **<Add** button.

Be sure to click **Save** before closing Flex Configurator, otherwise changes are not saved.

### Set Up Taskmaster Flex to Process Your Document Images

You can exercise the power of Taskmaster Flex to process your own document images.

#### Delete Sample Input Images

You have backup copies of the four sample Taskmaster Flex images in the `c:\Datacap\Flex\images\Single Page` folder, as well as in the backup copy of the `c:\Datacap\Flex` folder that you stored outside of the Datacap folder, so you can delete the copies of the sample images in:

c:\Datacap\Flex\images\Input

### Copy Your Images to Input Folder

Now copy your own sample images to the c:\Datacap\Flex\images\Input folder.

### Define New Document Classes and Index Fields

You must now define any new Document Class(es) and the index entries you want for each class. You can use the existing Document Classes and indexes if they are appropriate.

To add a new Document Class, for that class, decide which fields you want as index fields, which are required, how each is formatted, and whether or not data validations are to be performed.

The format of the field and the type of data validations to be performed should match the format and validations used in the downstream system that receives the images and index data. That system is likely to be a document imaging application or an accounts payable application in an ERP system. You might contact the IT support personnel responsible for the receiving application to most accurately define the format and validations.

**Note:** For demonstration or evaluation purposes, consider adding only minimal data validation. This enables you to prove Taskmaster Flex's ability to identify documents and fields without the effort required for defining all the necessary validations.

You also need to consider the sequence in which the index fields are to appear on the Verify panel. In general they should match the order that they appear on the actual document, though they do not have to.

Here are the necessary steps to add a new Document Class using the Flex Configurator:

- 1 Start Flex Configurator.
- 2 On the Flex Configurator Index Fields tab, add each of your new index field names, format and any validations. For more details, see **Flex Configurator Field Definitions**.
- 3 Click the **Save** button.
- 4 Click the Document Class tab and click the **Add** button below the list of Document Classes. Enter your new document name (class).
- 5 Make sure your document name is selected and then click first index name to be added and click the **<Add** button.
- 6 Add the remaining indexes for your new document.
- 7 Click the **Save** button.
- 8 Repeat from step 4 to add another Document Class and index fields.
- 9 Be sure to click **Save** before exiting Flex Configurator. The new Document Class or classes are now available in the Document Classes field on the Taskmaster Flex Verify panel.
- 10 Run Taskmaster Flex and test with your images.

**Flex Configurator Field Definitions**

<b>Form Item</b>	<b>Purpose</b>	<b>Generates Validation Error<sup>1</sup></b>
Index Field Name	<p>The name of the field or index. You can specify any alphanumeric character and spaces.</p> <p>Add new index field: Provide a new field name and any parameters to be specified and click <b>Save</b>.</p> <p>Modify existing index field: Click the index name, make changes to the parameters and click <b>Save</b>.</p> <p>Delete an index field: Click the index name and then click <b>Delete</b>. Warning: You are not prompted to confirm the delete, it takes place immediately.</p> <p>Limits:</p> <p>16 index fields per document.</p> <p>Maximum of 50 characters for a field name.</p>	N/A
Data Type	Values can be alphanumeric, numeric, currency or date. Currency is US \$. Any date format is acceptable, the date is verified using Microsoft date conversion routines.	Yes
Min Length	The minimum number of characters that can be entered. If left blank, any length is accepted.	Yes
Max Length	The maximum number of characters that can be entered. If left blank, any length is accepted.	Yes
Character Filter	The types of characters allowed in this field - all other characters are removed or “filtered” out. For example “1234567890-ABC” limits the field to only those 14 characters. To filter out dashes, ‘-’ in a social security number use the filter “0123456789.”	No

<sup>1</sup>The error appears on the Verify panel when the data in the field does not meet the defined validation criteria.

Form Item	Purpose	Generates Validation Error
Picture String	<p>A = Alphabetic characters only. To allow numeric and/or punctuation characters in an Alpha field, specify a picture string other than A.</p> <p>a = Alphabetic and punctuation characters.</p> <p>D = Dates. The dates must be expressed with numeric characters. You can delimit months, dates, and years with hyphens, periods, and forward slashes.</p> <p>F = Float or fractional numbers. A period for decimal separation is permitted as is a minus sign for negative numbers.</p> <p>f = Whole numbers and punctuation marks.</p> <p>L = Lower-case alphabetic characters only.</p> <p>l = Lower-case alphabetic characters and punctuation characters.</p> <p>N = Whole numbers only.</p> <p>P = Punctuation characters only.</p> <p>T = Time values. These values are expressed in numbers. In addition, the characters P, M, and A are allowed to distinguish between morning and afternoon times, and colon characters are allowed to delimit hours, minutes, and seconds.</p> <p>U = Upper-case alphabetic characters only.</p> <p>u = Upper-case alphabetic characters and punctuation characters.</p> <p>X = Alphabetic and numeric characters.</p> <p>x = Alphabetic, numeric and punctuation characters.</p> <p>Z = Allows any printable ASCII character.</p> <p># = Whole numbers and the minus sign.</p> <p>0 = Must be a dash “-.”</p>	Yes
Required	Is a value in this field required? Either Yes or No.	Yes

Form Item	Purpose	Generates Validation Error
Find by Zone	<p>Either Yes or No. The default is Yes. This determines whether field values are searched for by a zone on the page. The default is Yes and it normally should remain so. The first time the document is processed by Taskmaster Flex, you define the zone by using Click N Key or drawing a box around the data on the image. If the position of the value is highly variable, (for example a total line), then use No. When No is specified, the value can be located either by using a regular expression or general locate searching. Note: This parameter controls the locating of a value after it is found or identified on the Verify panel the first time.</p>	
Locate Method	<p>This indicates the method by which the field is found when the <b>Locate</b> button is clicked the first time the document is displayed on the Verify panel. Valid values are Keyword File, Regular Expression, or None.</p> <p>Keyword File: Indicates a file contains the keywords to search for. Taskmaster Flex searches for the keywords and then searches the area around the keyword to find the data.</p> <p>Regular Expression: Indicates a pattern is used when searching to identify strings of text.</p> <p>None: No attempt is made to automatically locate the index value. On the Verify panel, you must identify the index by clicking it, or clicking and dragging the mouse to draw a box around it.</p>	
Keyword File	<p>Identifies the name of the file containing the list of keywords used to locate the index file. It must be located in the \Flex\Process folder. There is a separate keyword file for each index field. The keyword file must have an extension of .key. The following keyword files are provided by default: InvDate, InvNum, IsDate, PONumber and TotalDue.</p> <p>As an example, the default PONumber.key file contains:</p> <p>Purchase Order #:  P.O. #:  P O #:  PO#  ORDER #:  P.O. NUMBER  Your P.O. No</p>	

Form Item	Purpose	Generates Validation Error
Regular Expression	<p>Identifies the regular expression to be used when searching. A regular expression, sometimes called a pattern, provides a concise and flexible means for identifying strings of text of interest, such as particular characters, words, or patterns of characters. Here is the default regular expression used for the social security number (SSN):</p> <pre data-bbox="548 598 1071 661">[^\b\s\n\r][0-9]{3}[-]{1}[0-9]{2}[-]{1}[0-9]{4}[\b\s]*</pre> <p>Here is the regular expression used for the date data type:</p> <pre data-bbox="548 766 1153 829">[^\b\s\n\r]([0]?[1-9] [1][0-2])[./-]([0]?[1-9] [1 2][0-9] [3][0 1])[./-]([0-9]{4} [0-9]{2})[\b\s]*</pre> <p>Note: The format for regular expressions is not part of this documentation. There are many sources of information about regular expressions on the Internet. Here is a link to the Wikipedia definition:</p> <p><a href="http://en.wikipedia.org/wiki/Regular_expression">http://en.wikipedia.org/wiki/Regular_expression</a></p>	
Refresh Button	The <b>Refresh</b> button is used to re-establish a lost database connection.	

# Taskmaster FlexID

## *Introduction to Taskmaster FlexID*

One of the first steps to the document identification process is to separate the scanned pages into individual documents. For example, an email may have fifteen images attached to it. Do those images represent fifteen separate documents? Or do they represent two documents of six pages and another three page document? Of course when you view the images you can see where each document begins, the number of pages in the document, even attachments to a document.

Taskmaster FlexID provides simple page identification without the use of separator pages. It displays thumbnails of document pages and enables you to specify when a page is the beginning of a new document, the continuation of the previous document, or an attachment to the previous document. In addition, it allows you to re-order pages should they be presented in an out-of-order sequence.

Taskmaster FlexID features include:

- **No document setup.** Just point Taskmaster FlexID to your document page images and you are ready to begin specifying the type of each page.
- **Works with any document.** You are not limited to invoices or any other pre-determined document type. Any electronic document can be used with FlexID and FlexID can be used with any Taskmaster application.
- **Flexible user interface.** Page types can be set by a mouse using a pull down menu or by a hotkey associated with each user-definable page type.
- **One-file configuration.** A single settings file defines the color characteristics of the user interface, pages types, and short-cut hotkeys.
- **Simple reordering.** Just hold down the Shift key, click a document, and nudge it left or right to move the page. Fast and simple.

## *Using Taskmaster FlexID*

Taskmaster FlexID is a component of Taskmaster Flex. It is accessed from the Flex Client.

While the sample images used in the Taskmaster Flex *Demo Job* are images installed as part of the Taskmaster Flex installation process, the sample images mentioned here as part of the Taskmaster FlexID *Demo-FlexID Job* are hypothetical images meant to give you an idea of the steps required to adapt the Taskmaster Flex *Demo-FlexID Job* to process your own images.

### **Multi-Page TIFFs**

The hypothetical images processed by the Taskmaster FlexID *Demo-FlexID Job* are multi-page TIFFs containing three different medical records, each of which contain 2 to 3 pages for a total of 8 pages/images. The documents are: History and Physical, Patient Assessment and Summary of Laboratory Tests.

### Setup Document Classes and Index Fields in Flex Configurator

Use Flex Configurator to define the following three Document Classes (types) and three index fields:

Class	Fields	Format / Validations
History and Physical	Patient	Alpha
(3 pages)	Medical Record Number	Alpha, filter: M0123456789, picture string: ANNNNNNNNN
	Report Date	Date
Patient Assessment	Patient	Alpha
(3 pages)	Medical Record Number	Alpha, filter: M0123456789, picture string: ANNNNNNNNN
	Report Date	Date
Summary of Lab Tests	Patient	Alpha
(2 pages)	Medical Record Number	Alpha, filter: M0123456789, picture string: ANNNNNNNNN

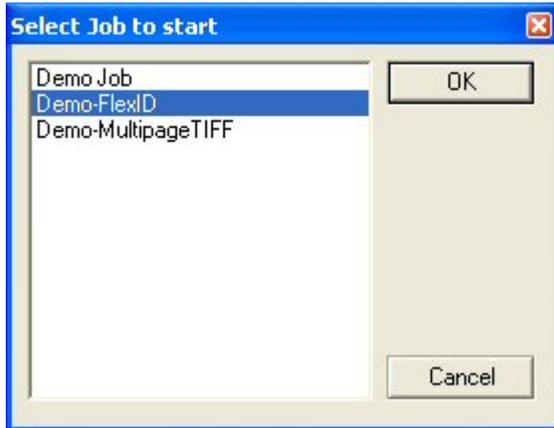
### Tasks Performed by Demo-FlexID Job

The *Demo-FlexID Job* differs from the *Demo Job*. It has five tasks, one of which is the FlexID task:

- **VScan** – This virtual scan task brings the sample document images into Taskmaster FlexID.
- **FlexID** – This task displays thumbnails of document pages and enables you to specify when a page is the beginning of a new document, the continuation of the previous document, or an attachment to the previous document. In addition, it allows you to re-order pages when they are presented in an out-of-order sequence.
- **Rulerunner** – This task performs many image processing functions including image cleaning (e.g. despeckling), image rotation (if necessary), full page character recognition and more. It also attempts to identify the document comparing it to a library of known document types.
- **Verify** – This task displays each separated document in the Verify panel. If the document is not known to Flex, you identify the document type by selecting it from a dropdown list, and then identifying where the data fields are located. If the document is known to Flex, the Verify panel displays the fields and data so you can ensure they are correct.
- **Export** – This task saves information about the documents and exports field data to XML files for input into a back-end application such as a document imaging or ERP system.

### Run VScan Task

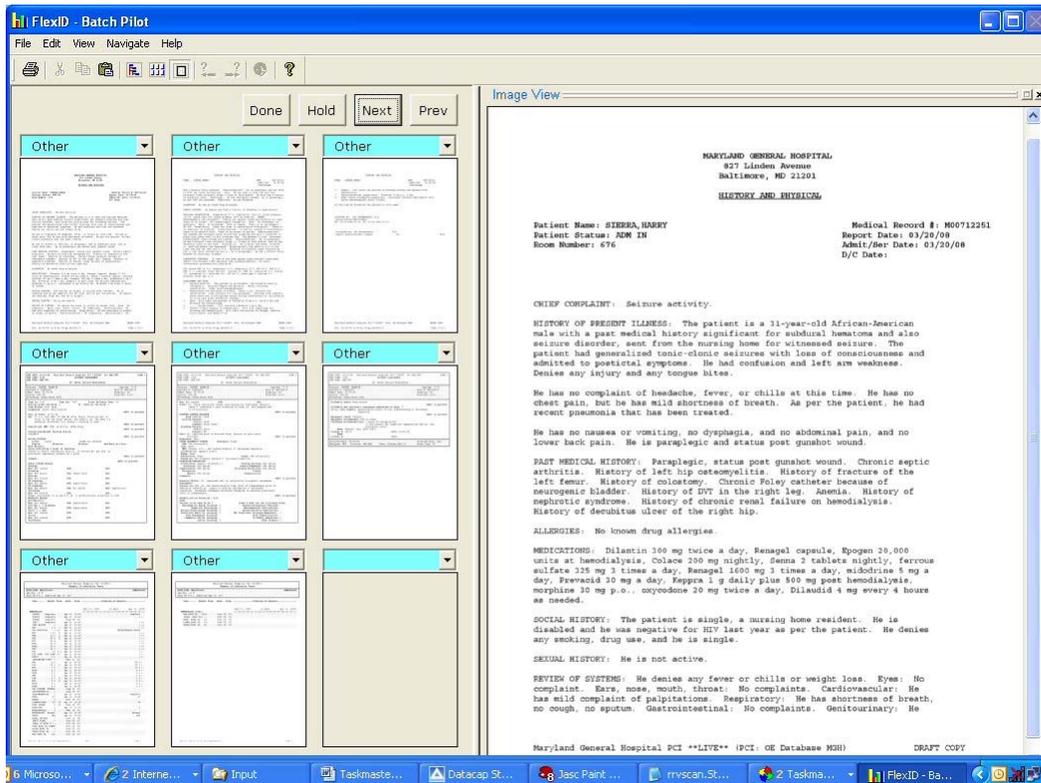
Double-click the **Scan** icon, select *Demo-FlexID* and click **OK**.



A Rulerunner status window briefly appears as the images are processed. When processing is complete, a message box is displayed indicating the task has finished. Click **OK**. Another message box asks “Do you want to continue processing?” Click **No**.

### Run FlexID Task

Double-click the FlexID icon. This task displays thumbnails of document pages:



FlexID enables you to visually and quickly identify each page by selecting the appropriate type of page from the drop down field above each page:

- *Main\_Page*: The beginning of a new document.
- *Trailing\_Page*: A continuation page of the current document.

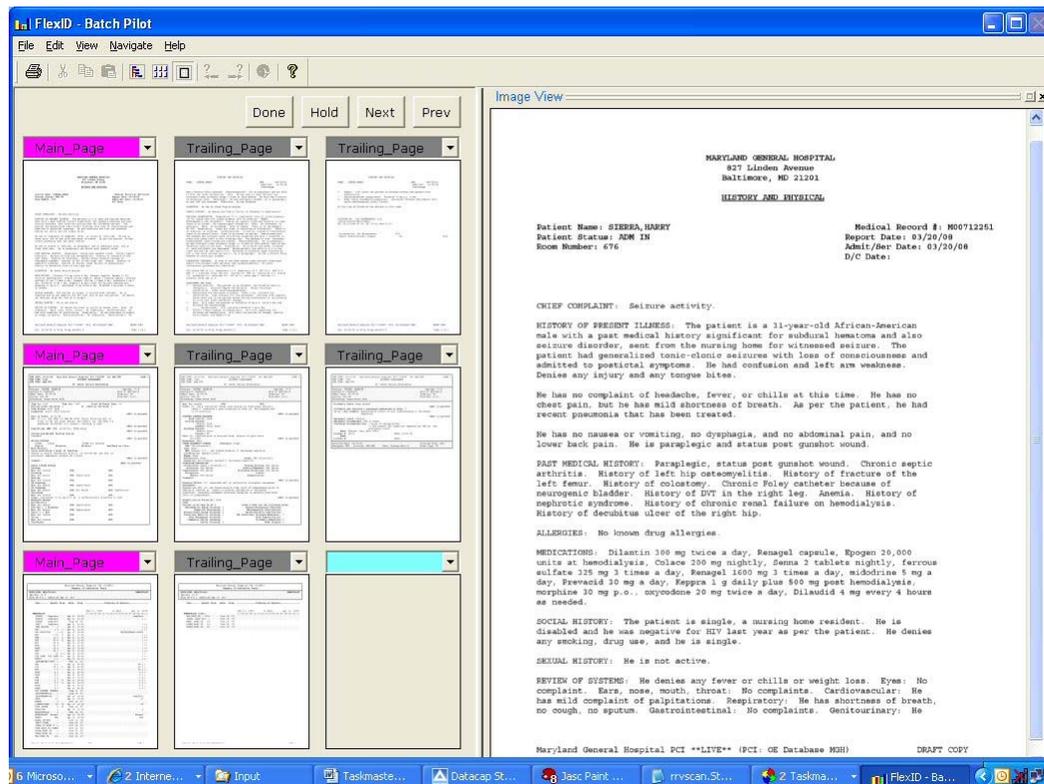
- **Attachment:** An attachment to the current document. OCR is not performed on attachments, therefore attachments cannot have indexes.

As you select the page type for each page or if you click the thumbnail image, the large image reflects the current selection. Go through and identify the separate documents contained in the stream of pages.

### Note on Page Re-Ordering

The pages can be re-ordered one page at a time either to the left or to the right. To re-order, hold down the Shift key and use the mouse pointer to drag the thumbnail image, or just nudge the image to the right or left.

Here we see the identified pages:



Click **Done** when complete.

### Run Rulerunner Task

Double-click the **Rulerunner** icon or double-click the batch number (e.g. “1”) in the Job Monitor window.

The Rulerunner status window is displayed as the document images are processed. This may take a few minutes, as Rulerunner is running a number of processing-intensive modules, including PageID, Recog, field Locate, and others.

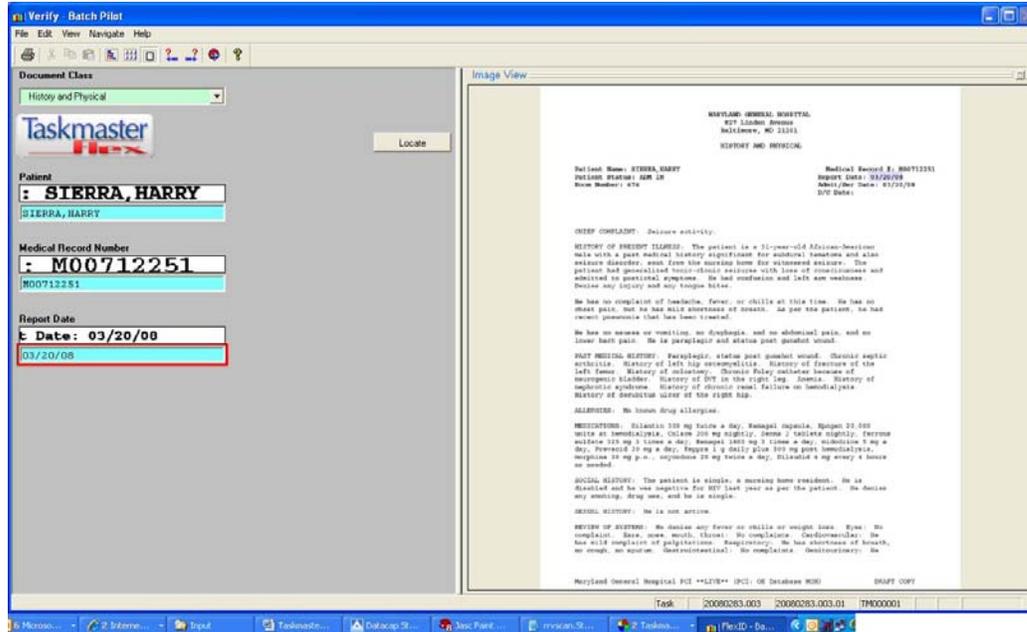
### Run Verify Task

When the Rulerunner task completes, refresh the Job Monitor window, and run the Verify task by double-clicking the **Verify** icon in the Operations window, or double-clicking the batch number of the batch you want to run in the Job Monitor window.

## Quick Start Guide

The Verify task displays the document image and the **Document Class** field from which you select the type of document.

Below we see the Verify panel with the first image displayed, a History and Physical report. The **Document Class** has been selected and the data for the fields identified using Taskmaster Click N Key:



### Note on Multipage TIFFs

When a multi-page TIFF is displayed in the Verify panel, the first page is displayed. To view the next page within a multi-page TIFF, press SHIFT+ALT+PGDN. To return to the previous page within a multi-page TIFF, press SHIFT+ALT+PGUP. To view the next

document, click the **Next**  arrow button.

This image contains three pages. Obtain the necessary index data from the first page. When done, click the **Next**  arrow button to go to the next document in the batch, which is a Patient Assessment document.

Select the **Document Class**, use Click N Key to identify the fields, and click the **Next**  arrow button to continue to the next document.

Select the **Document Class**, and use Click N Key to identify the fields.

Exit the Verify panel.

### Run Export Task

When the Verify task completes, refresh the Job Monitor window, and run the Export task by double-clicking the **Export** icon in the Operations window, or double-clicking the batch number of the batch you want to run in the Job Monitor window. Export saves information about the documents and their fields, creates Taskmaster fingerprints, and exports field data to XML files.

When Export completes, refresh the Job Monitor window. The status for this batch of documents is now *Job Done*.

### Re-run Demo-FlexID Job to Show Auto-Classification

To verify that Taskmaster Flex now understands these medical documents, re-run the *Demo-FlexID* Job. Double-click the **Scan** icon, select *Demo-FlexID* and click **OK**. When the VScan task is complete, refresh the Job Monitor window, and double-click the job number for this new batch to run the FlexID task. Identify the image pages. Run the Rulerunner task. Rulerunner compares these documents against the library of known documents and finds matches. Taskmaster Flex now knows the **Document Class** and where to find the data for the required fields.

Refresh the Job Monitor window, and double-click the batch number to run Verify. Notice that Taskmaster Flex now displays the **Document Class** and the fields. The data for the fields is also displayed.

Go through and view each document by pressing the **Next**  arrow button. Taskmaster Flex has “learned” these documents and can now automatically determine the **Document Class** and extract the required field data. You do not need to run the Export task. Exit when complete.

### Deleting Taskmaster Fingerprints

Fingerprints are used for document identification. They are automatically created by Taskmaster Flex during the Rulerunner task. Normally you would let Taskmaster manage the fingerprints. However, there are times that you may wish to delete existing fingerprints. Deleting fingerprints between demonstrations “cleans up” and enables you to show selecting the **Document Class** and using Click N Key to identify the data fields on the documents. If you do not delete the fingerprints and then you give your demonstration using the same documents as the previous demo, existing fingerprints will not let you demonstrate Taskmaster Flex’s “learning ability.”.

### Datacap Studio

Datacap Studio is the application development tool for Taskmaster and you can use it to delete fingerprints.

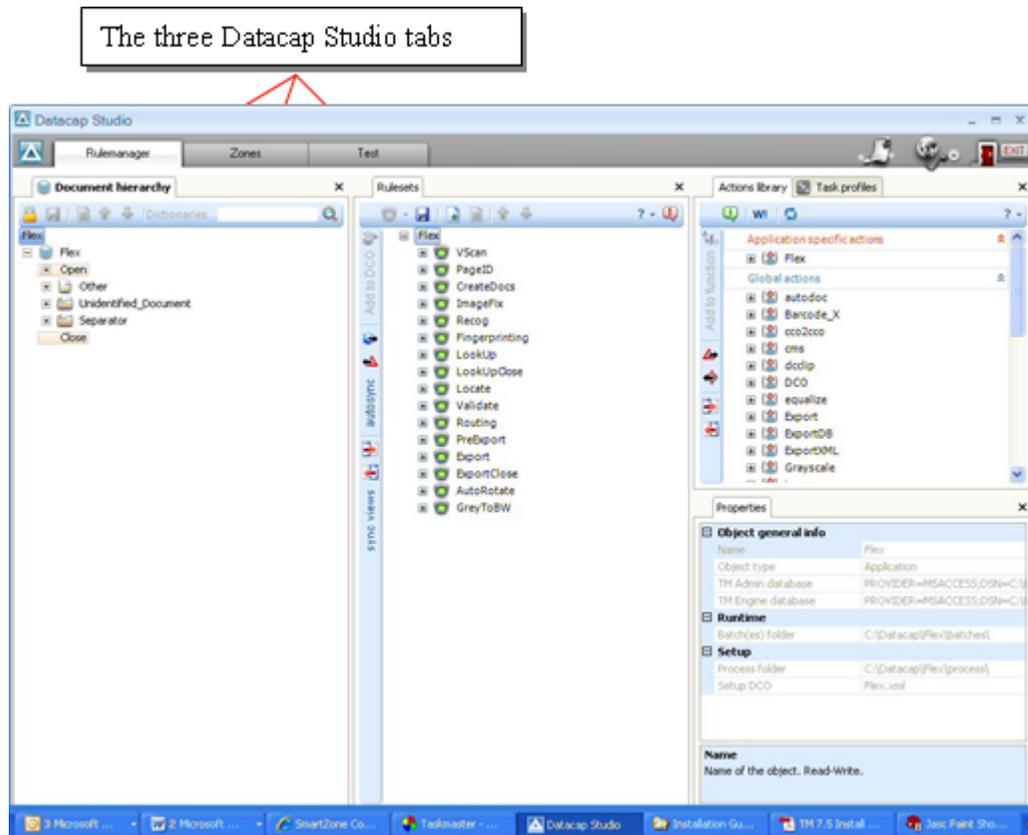
To start Datacap Studio:

- Ensure the Taskmaster Server Service is started.
- From your Windows **Start** menu, select **All Programs**, select **Datacap**, select **Datacap Studio**, and select **Datacap Studio** again. A Taskmaster Server dialog opens.
- Enter or ensure the **Server name or address** is **127.0.0.1**, then click **Next**. A Connection progress bar is displayed briefly, then the Select Application – Rulerunner Service dialog opens.
- Select the **Work offline. Internal Rulerunner service** option, then click **Next**. The Select Application – Applications dialog opens.
- Select the Flex application and click **Next**. A Connection Progress window is displayed briefly, then the Select Application – Taskmaster Login dialog opens.
- Select the **Taskmaster authentication** option, and ensure **User ID** and **Password** are **admin**, and ensure the **Station** is set to **1**, then click **Finish**. The Datacap Studio main window opens with the Rulemanager tab displayed on top.

## Quick Start Guide

---

Here is Datacap Studio; what is displayed to you in Datacap Studio may vary slightly:

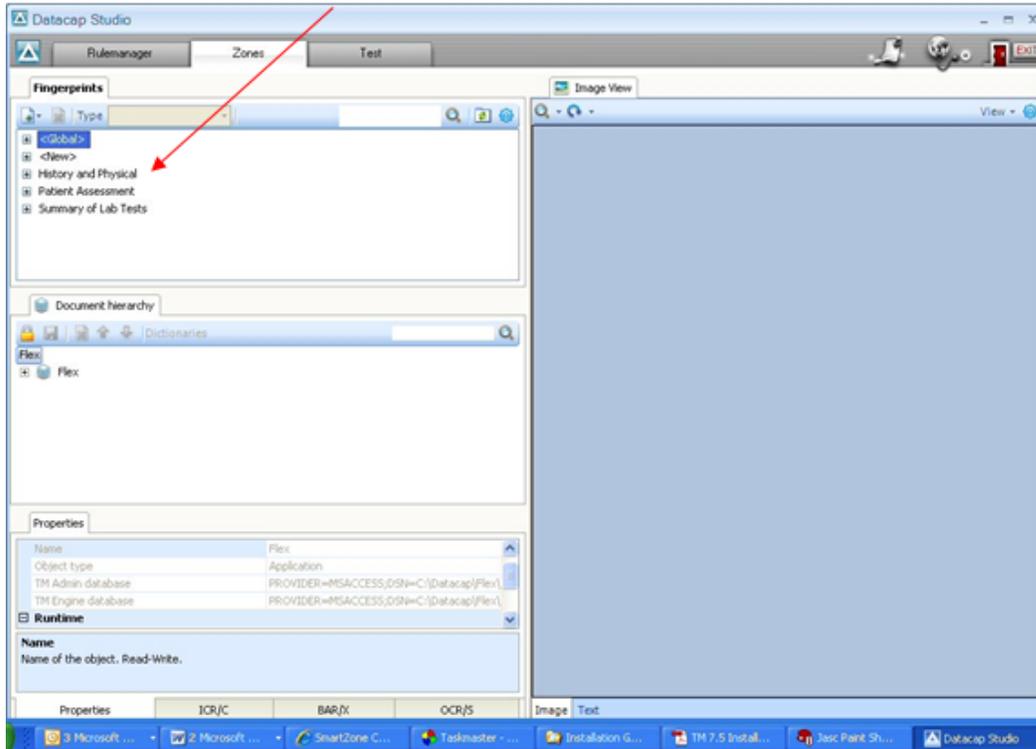


The Rulemanager tab is used to create and maintain Taskmaster applications. The Zones tab is used to view and maintain fingerprints and zones. (Zones are areas of the document containing data for the fields.) The Test tab is used to execute and test Taskmaster applications.

## The Zones Tab

Click the Zones tab. What is displayed to you may vary slightly from this example:

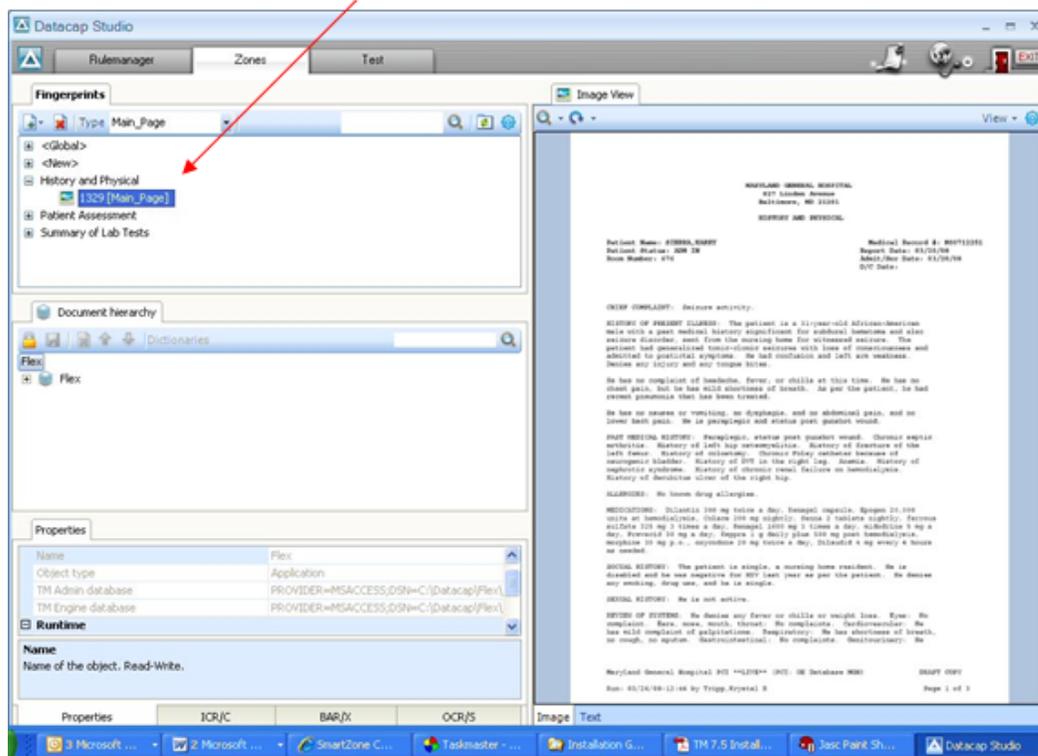
The defined document classes



The Fingerprints tab displays a hierarchical listing of the Taskmaster Flex Document Classes and the fingerprints that exist for each Document Class. The top of the hierarchy is the Document Class name. All the classes for which fingerprints were created are listed. We have three fingerprint classes in the above example: History and Physical, Patient Assessment and Summary of Lab Tests. We have these classes because we used Taskmaster FlexID with the three sample images described in this documentation. The <Global> and <New> entries are always present.

Click the plus sign (+) for the History and Physical Document Class and then click “1329 (Main\_Page)”:

A fingerprint for the “History and Physical” document class



The fingerprint for the History and Physical document is named “1329 (Main\_Page)”. When the fingerprint is clicked, the image is also displayed in the Image View area. We can have multiple fingerprints for a Document Class. For example, when you process five different History and Physical documents with Taskmaster Flex, the result is five fingerprints.

### Deleting Fingerprints

To delete a fingerprint, click the fingerprint to select it, and then click the **Remove Selected** button:

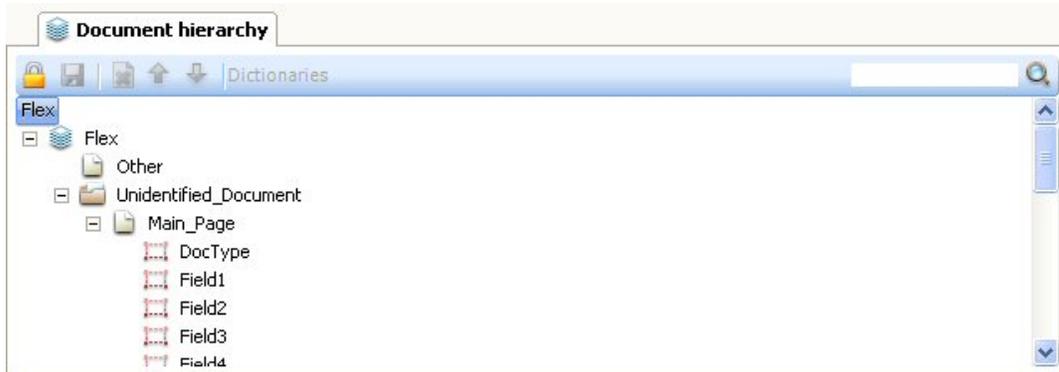
The fingerprint is immediately deleted, and the image is no longer displayed.

If your goal is to be able to present a Taskmaster Flex demonstration showing how new documents are processed and automatically identified, delete all the fingerprints you create with your documents between demonstrations. In our example above we would also delete Patient Assessment and Summary of Lab Tests.

### Viewing Document Zones

You can easily view the zones that have been defined for a document. First select a Document Class and fingerprint to display the document in the Image View area.

Under the Document Hierarchy tab, expand the hierarchy for the Flex application by clicking the '+' box. Then expand "Unidentified\_Document" and "Main\_Page" so you see the following:

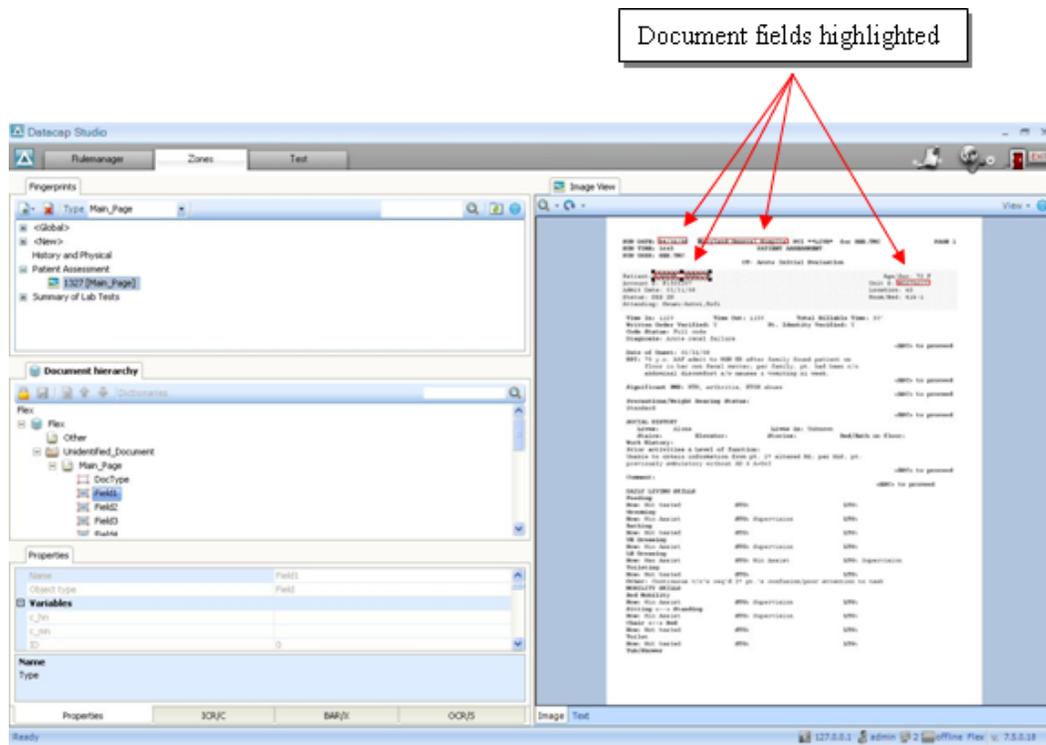


Now click **DocType** under **Main\_Page** and all the defined fields are displayed on the document image. You can also click a field name, e.g. **Field1**, to highlight that particular field.

**Note about Flex Field Names**

Taskmaster Flex field names are generic because Taskmaster Flex can be used with any document. The fields names that you have provided via the Flex Configurator tool are not displayed in Datacap Studio.

Here we see the Patient Assessment document image with its field zones highlighted and **Field1** selected:



Exit Datacap Studio by clicking the **Exit Door**  button.

