Associated Import Utility

Addendum to halFILE for Windows?
Document Storage & Retrieval

Version 2.1
January 26, 2004

hal

Systems Corporation

TABLE OF CONTENTS

Overview

The Associated Import Program, ASSOCIMP32.EXE, can be used to:

1)   Import files that are not standard halFILE Image Files (PDF files, Word documents, etc.) into the halFILE
      System to be stored with halFILE and viewed using the original program.
2)   Import Multi-Page TIFF images, and convert them into standard halFILE Image Files
3)   Fill out designated index fields by reading values from the name of the file being imported.
4)   Automatically execute multiple import filters with one click.

This document describes how the program works and how you can set you your import.

Requirements

This program requires halFILE version 2.1 or later and the AssocImp32 program. AssocImp32 should be installed with the halFILE for Windows 2.1 Setup program, and should be located in the same folder as the main halFILE Programs.

Running the program

To launch the AssocImp32 program, open up your halFILE program, login if security is enabled, and select Tools -Associated Import from the halFILE Manager Toolbar.

Operation

When you run AssocImp32, you are first presented with a screen where you select the Application, Database, Basket and Import Filter as shown below. The first time the program is run, no Import Filters will be listed.
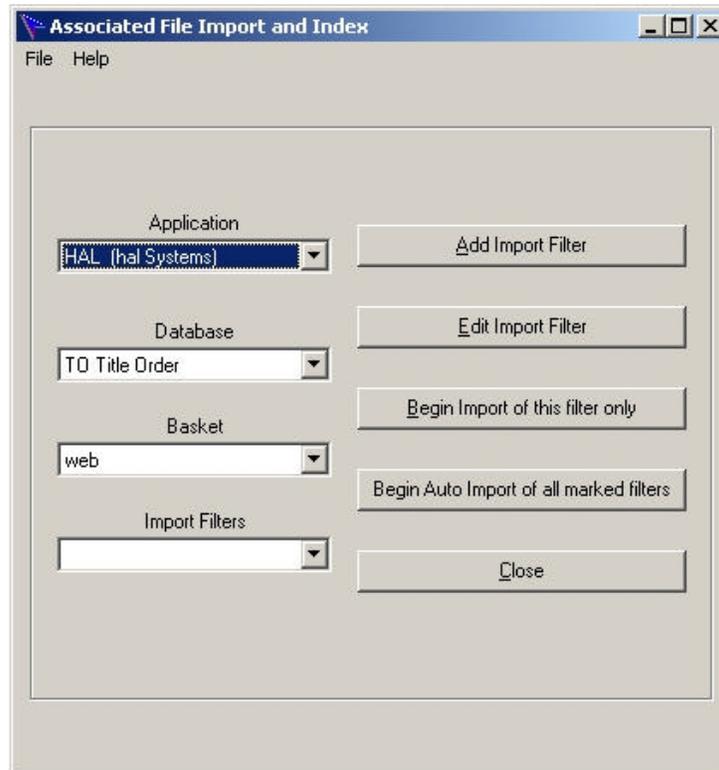
The following buttons are available on the Main Menu Screen:

<u>Add Import Filter</u> – this adds a new import filter for the selected Application and Database.

<u>Edit Import Filter</u> – this edits the import filter that has been selected in the Import Filter box.

<u>Begin Import of this filter only</u> – this executes the selected import filter, importing files into the system and indexing them as defined by the filter.

<u>Begin Auto Import of all marked filters</u> – this executes every import filter that is marked for auto import.

<u>Close</u> – Close the Associated Import program.

Adding a New Import Filter

When you click ADD Import Filter, the following screen is shown.
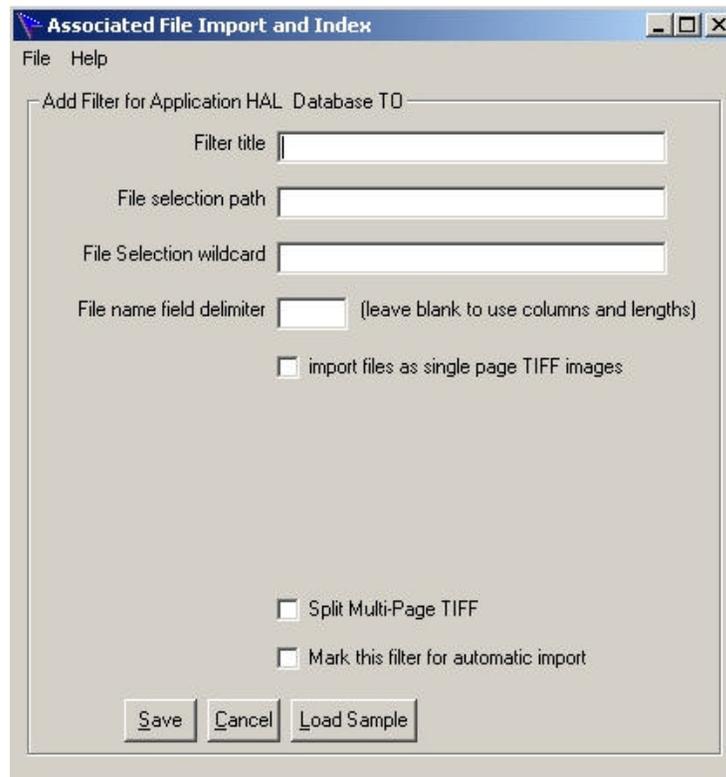
Filter Title – Free-form text used to describe the filter such as "Tax Statements"

File Selection path – this is the path where the files to import will be found. Be sure to include a trailing backslash (\).

File Selection wildcard – this is the selection criteria for the files to be imported. You may use the wildcard characters asterisk (*) and question-mark (?) in this field. For instance:

1) If you want all of the DOC files found in the designated path, you would enter '*.DOC'.
2) If you wanted only the PDF files that began with '264,' you would enter '264*.PDF'.
3) If you wanted only TIF image files that began with a 4-digit year, followed by exactly 4 additional characters, you would enter 2004????.TIF

Use the LOAD SAMPLE button to fill in the File Selection boxes by browsing the file system.

File name field delimiter – this designated the delimiter (or character) that is used to identify fields within the file name. If you leave the field delimiter blank, the system assumes that you are not using a delimiter; rather you will designate the starting position and length for each field in the file name.

Import files as single page TIFF images – check this box only if the files being imported are single page TIFF images. This will enable the Append Page options (See *Editing an existing Import Filter, page 6*)

Split Multi-Page TIFF – use this option if the files you are importing are Multi-Page TIFF files that need to be split into single-page TIFF files before importing into halFILE. Selecting this option will automatically select the Single Page TIFF check box. If you are unsure of this option and its ramifications, please contact Technical Support.

Mark this filter for automatic import – indicate to the program that this filter will be included in an automatic import.

Once you have entered these options, Click the SAVE button to save these settings to the Filter.

Editing an existing Import Filter

When you click Edit Import Filter, the following screen is shown:
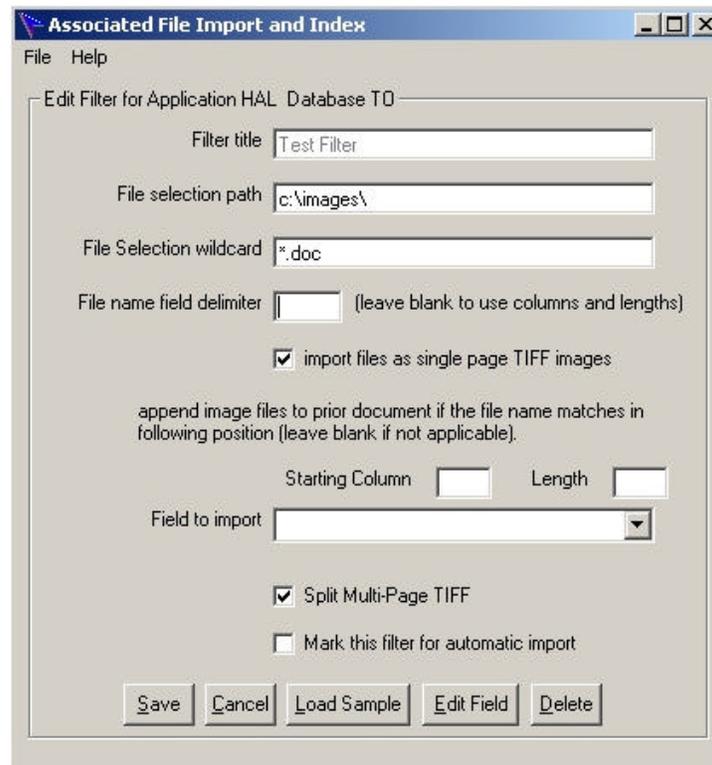
Import files as single page TIFF images - If you check this box, the append file options appear, indicating the Starting Column and length text boxes should be filled in. Filling in these boxes will allow the program to automatically staple pages together based on a certain string in the file name. For example:

1)   Images are single-page TIFF images.
2)   File Names are XXXXXXXX_PPP.tif (XXXXXXXX represents a file number, and the PPP is the page)
3)   Any docs that match the file number would be stapled together in halFILE with the same document number.
4)   You would put the number 1 in the starting column box, and the number 8 in the length box.

Field to import – when you are editing an Import Filter, you have this drop down box to define the fields that will be automatically indexed. You select the field from the drop down box, and then click the Edit Field button, which displays another screen (see *Editing the Field Information* section below).

You would select only those fields that need to be indexed when the Import Filter is processed.


The buttons at the bottom of the Edit Import Filter screen include:

Save	Saves the import filter, and returns to the Main Menu
Cancel	Cancels any changes you have made to the filter, and returns to the Main Menu
Load Sample	Lets you select a file name through the Windows Explorer interface, and then automatically fills in the File Selection path and File Selection wildcard boxes
Edit Field	Edits the selected field information
Delete	Deletes the current Import Filter (after confirmation), and returns to the Main Menu

Editing the Field Information

If you are using the Associated Import Utility to index documents based upon information found in the file name, selecting a field from the drop down, and clicking the EDIT FIELD button will take into the Field Posting options.

If you are using a delimiter to separate the field in the file name, you see the following screen:
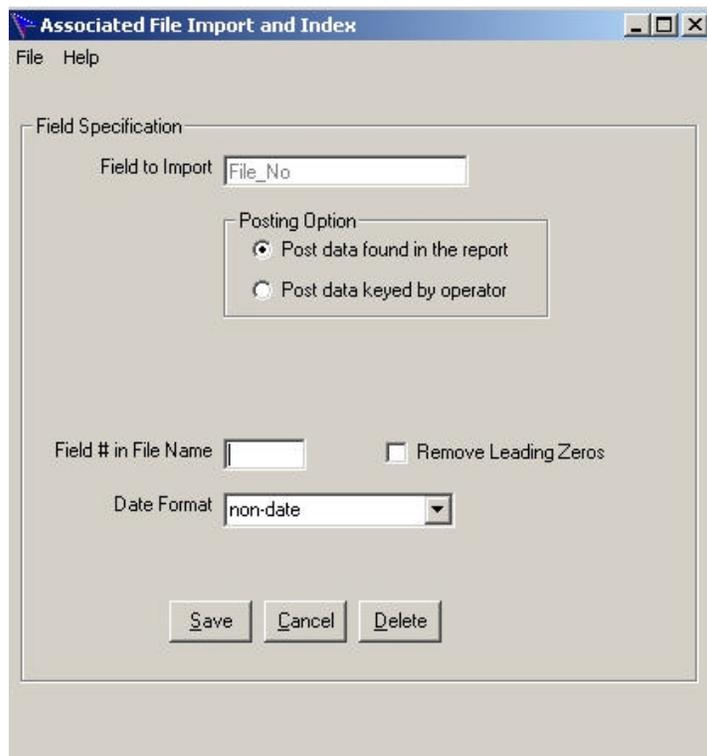
Posting Option: The 'Post data found in report' option is selected if you want the system to parse a part of the PDF file name and automatically fill in the index field. The 'Post data keyed by operator' option is selected if you want the system to ask you for the data to post when you run the import filter.

Field # in File Name : This option designates which field in the file name to post, counting the delimiters as separators:

Given a file name **DT_01262004_0738_0005.PDF**, the following fields could be read from the file name:

Field 1 DT   Instrument Type
Field 2 01262004   Date field
Field 3 0738   Book
Field 4 0005   Page

Remove Leading Zeros: Check this box, if you want extra zeros (0) removed from numeric fields. Using the previous example, Field 3 would become 738 instead of 0738, and Field 4 would become 5 instead of 0005. Be advised that removing leading zeros can cause record sorting issues in the halFILE Search Module Hit List and Crystal Reports (i.e., 1 would be followed by 10, followed by 100, etc).

Date format: This defines how the date is formatted in the file name. If the selected field is not a date field, leave the box as non-date, and continue. If the selected field is a date field, select the appropriate option from the drop down list. Using our example, '01262004' the date format would be 'mmddyyyy'. Note that to prevent errors during import, it is important to consistently use the same date format in all file names that will be used for a given filter. Also, you should zero-fill the month and day fields consistently (01031999 not 131999) to ensure that the system will correctly index the date field.

DO NOT check the Remove Zeros box if you are using a date field.

If you did NOT designate a field delimiter, the field information setup screen will be slightly different:

Posting Option – same as above

Select column – designates the starting column for the data to post.

Length – designates how many characters from the file name to use.

Remove Leading Zeros – same as above

Date format – same as above

If using the Starting Column/Length options, the example above would be set up like this:

Given a file name **DT_01262004_0738_0005.PDF**, the following fields could be read from the file name:

Start Column	Length	Value	
1	2	DT	Instrument Type
4	8	01262004	Date field
13	4	0738	Book
18	4	0005	Page

Executing an Import Filter Manually

Once your import filters have been set up, you can execute an Import Filter. To do this:

1) Open the Associated Import program
2) Select the Application, Database for the filter you wish to execute
3) Select the Filter you wish to execute
4) Select the Basket where the images will be placed
5) Click the button labeled **Begin Import of this filter only**
6) A list of the files located by the filter will appear
7) You will be asked three (3) questions:
   a. **Confirm execution of the filter**
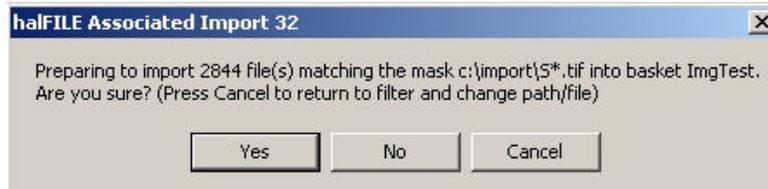
   1. If you answer YES, the import will proceed.
   2. If you answer NO, the import will stop, and you will be taken back to the main menu.
   3. If you CANCEL, you will be given the option to change the filter path and selection criteria.

   b. **Status in basket of imported docs**

   1. If you answer YES, the documents will be imported, indexed, and placed into the basket with the status of SCAN COMPLETE.
   2. If you answer NO, the documents will be imported, indexed, and placed into the basket with the status of INDEX COMPLETE

   c. **Deletion of original files**

   1. If you answer YES, the original files will be deleted after being copied into the basket folder.
   2. If you answer NO, the original files will be left intact after being copied. Always select NO if importing directly from a read-only media, such as a CD or DVD.

8) The program will begin the import, displaying a count of the number of files processed, and updating the list showing the file currently being processed.
9) Once it finishes, a button will appear labeled **DONE**. Click this to return to the Main Menu.

Executing Import Filters Automatically

Once your import filters have been set up, and you have tested the results, you can run these filters automatically. To do this:

1) Make sure the import filters you wish to run automatically have been set up with the **Mark this filter for automatic import** box checked (See **Adding a new Import Filter**).
2) Once confirmed, you can open the Associated Import program, and click the button labeled: **Begin Auto Import of all marked Filters**.
3) This will query the List of filters for any that match your selected Application and Database, and are marked for auto import, and display them on a list.
4) Check the box below if you want to delete the original files.
5) Click the button labeled **Begin Auto Import**
6) This will execute the filters listed sequentially

This process can also be run from a command line by using the command-line switch RUNAUTO. To do so, enter the command:

> <halFILEPath>\ASSOCIMP32.EXE RUNAUTO

Only execute this switch if you have fully tested all automatic import filters thoroughly.

Additional Options

halFILE Key Verify Support – beginning with version 2.1.4, the AssocImp32 Program will check to see if Key Verification is setup on the system, and if it is, documents that would be put in a basket as status INDEX COMPLETE will instead be placed in the basket as status INDEXED/NOT VERIFIED. This can be turned on and off by the following setting in the [VERIFY] section of the xxxVERyy.INI file, located in the halFILE Program Directory:

> VerifyAssociatedImport=[TRUE/FALSE]

SQL Transaction Log Dumping – beginning with version 2.1.4, the AssocImp32 program executes a DUMP TRANSACTION query against the Main database every 100 records. This is used to prevent the SQL transaction log from getting filled up, and also to slow down the rate of updates to the SQL Server, to prevent the SQL Server from locking up from excessive updates (what Microsoft refers to as a FIREHOSE operation). In some situations, if the number of docs imported by the filter is less than 1000, turning this option off can speed up the execution of the import filter. This option can be turned on and off by the following setting in the [DEFAULT] section of the ASSOCIMP32.INI file, located in the halFILE Program Directory:

> DumpTran=[TRUE/FALSE]

Troubleshooting

Log Options – the associated import program uses a log file where it writes process information, as well as additional information useful in debugging problems. The default log will be located in the TEMP directory, on the station from which the import filter was executed, and is called ASSOCIMP32.LOG.

If you have problems with setting up or running the Associated Import program, please contact halFILE Technical Support.