matchIT User Manual v5.2

Table of Contents

Important Notice

Preface

This is the User Manual for matchIT version 5.2. See the Getting Started Guide for an introduction to the main features of matchIT. For more in-depth information on some topics, refer to the Online Help.

Not all sections of this manual may apply to your copy of matchIT. If you would like further information about any of the modules you do not currently have, contact your supplier (see "Contact Us" on page 173).

Preliminaries

matchIT Configuration

Hardware Requirements

For running matchIT reliably, we recommend a PC with a minimum of a Pentium IV or equivalent chip with at least 512MB of RAM (memory).

See also: "Disk Space" below.

Activation

If you do not have an Activation floppy disk, you will need to contact your supplier for an activation code before you can use matchIT to process anything other than the supplied example data. An opportunity to enter a code is provided as part of the installation procedure, but it is also possible to run the activation program after installation, by selecting Password Activation from the Windows Start menu.

Display Properties

If you find that some of the dialog boxes are not fully visible after you have installed matchIT, make sure your screen resolution is set to at least 1024x768. In addition to this, the Color Palette should ideally be set to either "High Color (16 bit)" or "True Color (32 bit)".

Disk Space

You need free disk space of up to six times the size of data file(s) that you are processing (depending on what kind of processing you are doing) plus 100MB. There are ways of economizing on disk space, which your supplier can discuss with you in the light of your particular requirements.

If you reformat your disk, please read "Moving matchIT to another PC" in the Going Further with matchIT Guide first.

See also: "Hardware Requirements" above.

Operating System

matchIT runs under Windows 2000, XP, & Vista.

Installing

To install matchIT, please follow the instructions on the CD cover and in the Getting Started Guide. matchIT will guide you through the installation process. If this does not work, please call your supplier for guidance (see "Technical Support" on page 171).

New Features

matchIT version 5.2 has many new features available, including:

- Ability to work from a centralized file system (see Configuring Centralized Files in matchIT on page 160)

- Increased support for large Suppression files (see Support for Additional Suppression Files on page 75)

- Added multi-level matching functionality (see Multi Level Matching on page 101)

- Additional mailing exclusion tracking (see Non Mailable Records on page 46)

- Intelligent Data Merge enhancements (see Default Priority on page 92)

- Data archiving of a matchIT setup (see Data Archival of Jobs on page 16)

- Quality scoring on contact elements within a record (see Quality Scoring on page 144)

General Operating Techniques

Using Online Help

For Windows-style help while matchIT is running, access the Help menu. About matchIT will display which version of matchIT is installed. Using the Windows-style help, it is possible to use the contents or index tabs to find a topic and also search for a keyword.

See also:

- "Problems" on page 166
- "Further Information" on page 172

Basic Navigation

Menus

Menus follow Windows standard usage. All of the toolbar icons are available via the menus, with the exception of View Data which is also available as a hyperlink at the top of the Information Window.

Shortcut Keys

In the section which follows, the + symbol is used to mean pressing two keys at once; for example, Ctrl+End means "hold down the Ctrl key and press the End key".

Some functions do not require the mouse to find an option in the menu bar. *Shortcut keys* can provide a faster method of access.

The keys are listed below:

- **F1** – open the help file
- **F3** – close the current view and save changes
- **ESC** – close the current window without saving changes
- **CTRL+F** – find
- **CTRL+G** – find again
- **CTRL+C** – copy
- **CTRL+X** – cut
- **CTRL+V** – paste
- **CTRL+Y** – redo
- **CTRL+Q** – quit matchIT (from the main window)

Additional Menu Options

matchIT displays various menus which depend on the particular activity that is taking place. Some of the options on these menus are not always available – those that are not available appear as "grayed out" (fainter).

The Toolbars

*matchIT's **Main toolbar***
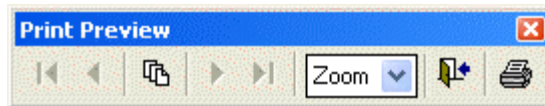
*matchIT's **Print Preview toolbar***

The toolbars allow you to select the most common menu options with one mouse click. If any of the icons are grayed out, either they are not available in your version of matchIT, or the layout of the Main File that you are using does not support that option.

To see what each icon does, hold your mouse pointer over that icon until popup text appears.

*Selecting a **Destination** of Printer gives more control over printing than the Printer icon in **Print Preview**.*

Selecting Print from the icon in the Preview toolbar prints one copy of the whole report to the default printer. If you want to print the report to a printer other than the default printer, print more than one copy or a page range, you should select a Destination of Printer from the relevant reporting options screen, rather than print from the Print Preview toolbar.

Main File

Most of the options from the matchIT main menu require a Main File to be selected. The Main File is a DBF file that usually contains name and address records in a structure prescribed by matchIT. If a Main File is currently selected, its name and directory appear at the top of the Information Window. If a Main File is not currently selected, a message to that effect appears there instead.

Open Main File

If a Main File is selected but you want to select a different Main File, select the Open option from the File menu or the "Open File" button. If a Main File isn't selected and you select an option that requires one, the Open Main File option will automatically be presented.

You then see a list of files to choose from, in the directory from which you last selected a Main File.

If the file has fewer records than "Max database size to display stats on info screen" (on page 13), when you open the Main File matchIT will count the number of flagged records, salutations and postcodes are displayed when the pointer is moved over the file's hyperlink, for display at the top of the screen – otherwise it displays "n/a" for these items.

Copying Main Files and Layouts

Copy Main File

You can select this option from the Tools menu to copy an existing File, including the records within it. You may wish to do this to have both a backup and a working copy of a Main File you have just created. To enter a name for your new File, click on the ▫ button to the right of the box that appears. After creating the new File, matchIT will open it as the Main File.

Copy Main File Layout

Import Records

You can select this option from the Tools menu to clone the layout of an existing Main File. This option allows you to create a new Main File with the same layout as an existing one (but containing no data). If no Main File is open, matchIT will ask you to select a Main File layout to copy, then to enter the name for the new file.

Data can then be imported into this new Main File by choosing "Import Records" from the Import menu.
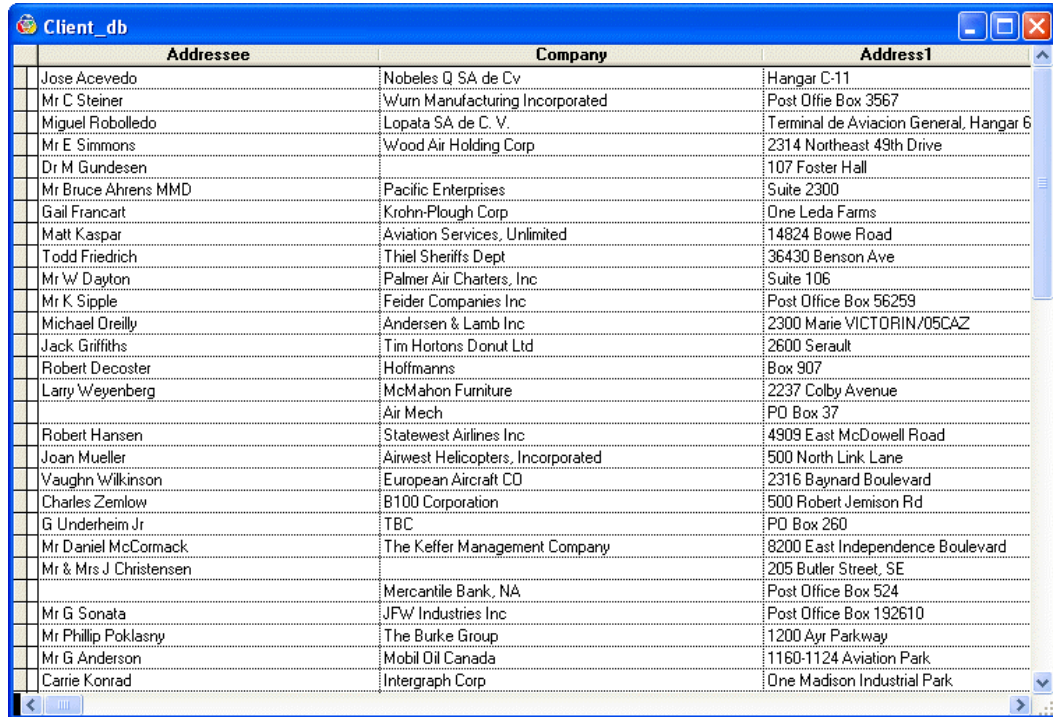
Closing Databases

matchIT automatically closes your currently selected Main File if any of the following are selected:

- "Close All" from the File menu
- "Open" from the File menu, unless selecting "Open Second File"
- "System Commands" or "Database Utilities" from the Tools menu

View Table

From several points in matchIT, you can view either your Main File (e.g. via the hyperlink button on the main screen) or other tables (from "**View Table**" under the **View** menu). This uses a Browse window.

The browse window displays the records in the selected table.

The field containing the cursor will be highlighted, (ready for any cut, copy, or delete operation that you might want to perform) after clicking in the window. **Tab** takes you to the start of the next field, and **Shift+Tab** to the start of the previous field.

When you have finished viewing (and possibly editing) the table, **Esc** closes the window but does not save changes to the field you are currently positioned on. **F3** or closing the window with the mouse saves changes to the current field as well as the rest of the table.

When you are browsing a table, a Table menu appears at the right hand end of the menu bar. This menu allows a variety of options for data manipulation, positioning within the table and for controlling the view of the table.

As with other Windows database applications, you can re-size columns and drag columns within the view, to position columns in the order that you want to view them – this does not change the underlying database structure. Alternatively, if you click on the thick black rectangle in the bottom left hand corner of the window, the cursor changes to a "split" icon which allows you to split the window into two panes, so that you can view widely separated columns simultaneously without having to drag the whole of a column over.

If you see a record that has been flagged but has not been physically removed ("packed"), it will have a thick black rectangle at the extreme left hand end of the row.

Performance Data

This is a menu option available from the View menu.

All Runs

During Import, matchIT writes a record of the key information concerning Import, matching, and deletion (including the Main File name) to PERFORM.DBF in the Database sub-directory. Selecting Performance Data, All Runs from the View menu allows you to view this table, with the most recent operation (i.e. "run") at the top.

Since Last Import

During these steps, matchIT also writes a copy of the progress information appearing in the Feedback Window to the PERFORM.TXT file in the same directory as the Main File. PERFORM.TXT can be viewed by selecting Performance Data > Since Last Import from the View menu. You must select a Main File first.

Setup Options

Introduction

matchIT's options have all been combined into one screen in version 5, which is accessible from the Jobs/Setup menu.

Operational Options
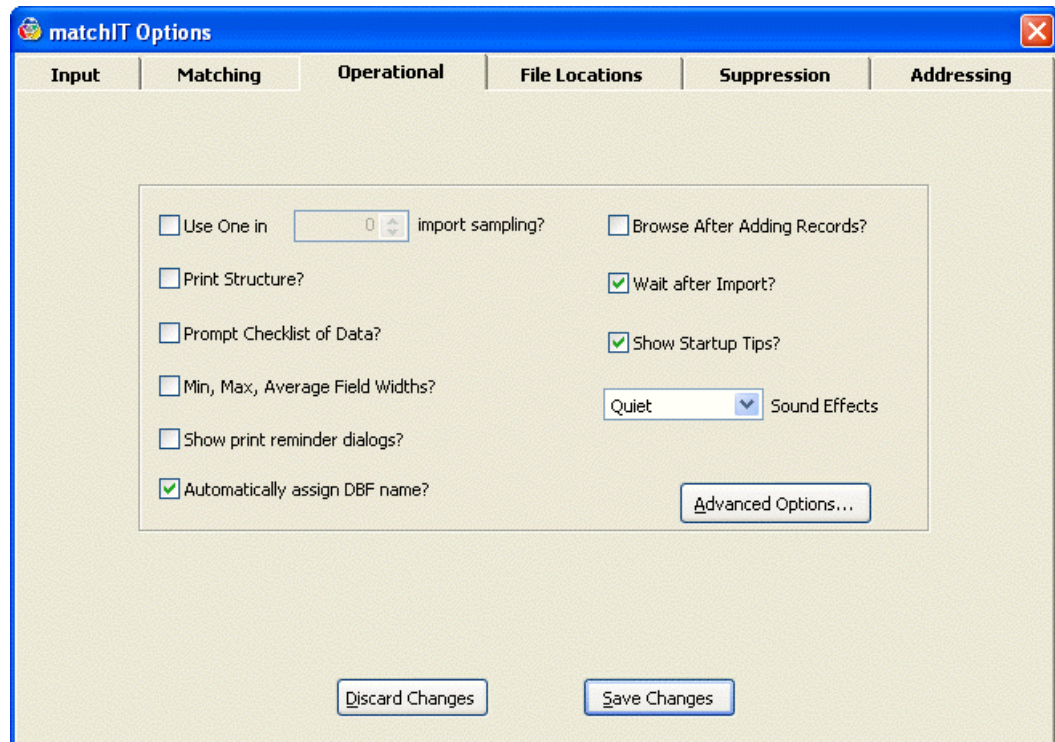
Use One in N import sampling?

Set this option ON to limit the import to an Nth sample of the input file. For example, set this option to 1000 to import only 1 in every 1000 records.

Print Structure?

Set this option ON to automatically display the Main File Layout report before Import. This report lists the different field names in the DBF file that you are importing into, as well as the widths and field types.

Prompt Checklist of Data?

Not available yet. Will be included in a future matchIT release.

Min, Max, Average Field Widths?

Set this option ON to automatically display the Main File Layout Field Widths report. This report lists everything that the Main File. Layout report does, but also displays the minimum, maximum and average field lengths within your records.

Show print reminder dialogs?

Set this option ON to prompt for printing reports after preview.

Automatically assign DBF name?

Set this option ON to automatically assign a name for the DBF file at the end of the Setup Wizard, in the same folder as the input file.

Browse After Adding Records?

Set this option ON to automatically view the contents of the Main File that your data has been imported into, before keys are generated.

Wait after Import?

If you tick the "Wait After Import" parameter check box matchIT will pause after Importing records so you can choose whether or not to proceed to Find Matches and if so which Match Keys to use. If you leave the box blank, matchIT carries on with matching automatically after Import, using default Match Keys.

Show Startup Tips?

When you start matchIT for the first time, you will see a startup tips screen with several options, one of which governs whether or not that screen will be displayed next time. You can set this parameter "On" so that you will see this screen when starting up, or "Off" to disable it.

Sound Effects

If you set the Sound Effects parameter to 'Off', matchIT will not make any beeps when displaying error messages or when it has finished importing, matching and outputting records. If you set it to 'Quiet', it will make a single beep for all occasions.

Advanced Operational Options
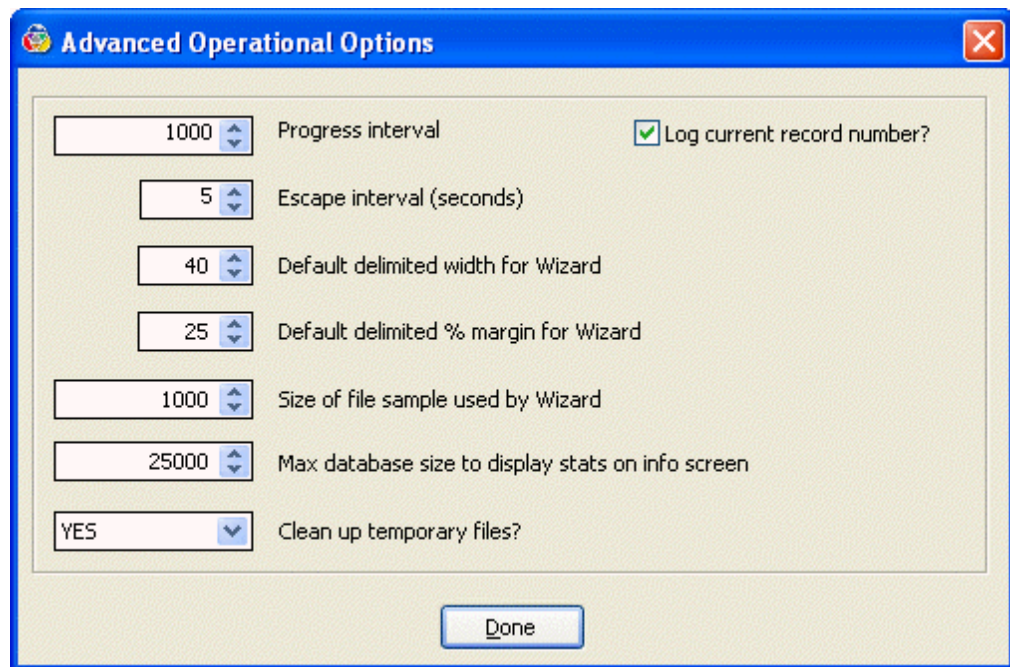
Progress interval

The Progress Interval parameter is an integer. Its value specifies the number of records to process between updates of the progress information on the screen. At the end of each progress interval, updates are written from a temporary buffer to the hard disk for the table being processed. Too low a progress interval slows processing down.

Escape interval

To maximize the speed of the importation and deduplication, matchIT only checks for any keyboard input from the user (to see if they have pressed escape to abort the process) at a set interval. This parameter is the size of that interval, in seconds.

Default delimited width for Wizard

When importing a file using the Setup Wizard, matchIT has to determine the widths of each field into which the data will be placed. If the file has a delimited format (Comma or Tab delimited), then matchIT assigns each field a default width. This parameter governs that default value. The exception is when the Intelligent Setup Wizard determines that a field is a postcode or ZIP, or when it appears to be of a fixed width in the file sample.

Default delimited % margin for Wizard

matchIT scans a sample of your data and determines the max length for each column, it then factors in a percentage to increase it by, to accommodate for any larger pieces of data that may not have been included in the sample. It's very similar to a margin of error when determining how matchIT generates field widths.

Size of file sample used by Wizard

When importing a file using the Setup Wizard, matchIT shows you a sample of your file to enable you to identify the contents of each field. This parameter governs the size of this sample; the sample is taken from the beginning of a file.

Max database size to display stats on info screen

matchIT displays details about the currently selected Main File on the info screen, such as the number of flagged records, and the percentage of records which have postcodes. If the file has a large number of records, however, these details can take a long time to calculate, making it inconvenient when opening files. This parameter sets a record limit above which these details are not automatically calculated when opening a new file. This option cannot be set to lower than 100 records.

Clean up temporary files

This parameter governs what matchIT does with temporary files when it closes down. If this parameter is set to 'Yes' then matchIT will always delete temporary files and if it is set to 'No' matchIT will always leave the temporary files. Alternatively, if this option is set to 'Ask', matchIT will check with the user before deleting temporary files when closing down. What constitutes a temporary file is defined in the table CLEANUP.DBF in the main matchIT directory, which contains a list of file masks such as "*.TMP"; you can add other masks to this list. matchIT will delete all files matching these file masks in the matchIT directory and all its installed sub-directories.

Log current record number

If you tick this checkbox, then matchIT will record the record number (or pair of record numbers and match key for matching) it is currently processing when it is importing a file or doing Find Matches/Overlap. The information is recorded in a text file, which is called IMPORT.LOG for importing, MATCHING.LOG for matching on a single file, and OVERLAP.LOG for a find overlap on two files; all these files are found in the main matchIT directory. This is useful when there is a problem during one of these operations, as you can inspect this file to see which record (or pair of records) was potentially causing a problem.

File Locations Options
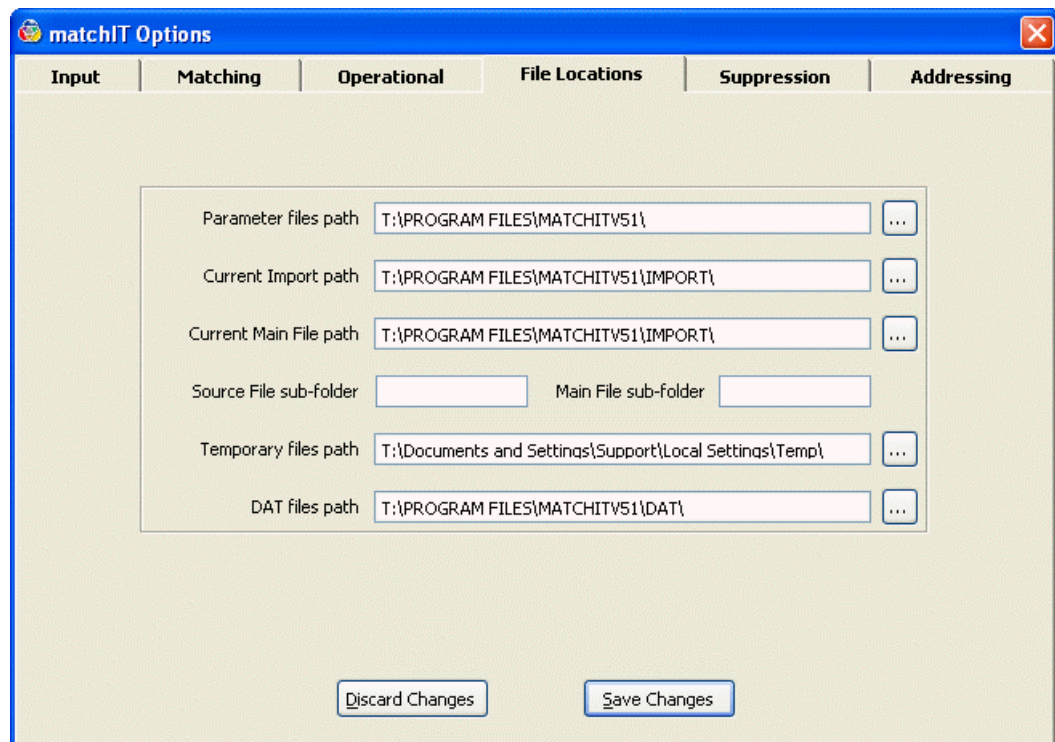
These options allow you to specify where matchIT should default to when looking for or creating various types of files.

Parameter files path

The Parameter Files Path from the matchIT Options, File Locations screen allows you to specify a directory for matchIT to look in for its parameter and matching results databases, instead of the default start directory and its sub directories. For example, if you want to use PARAMS and WEIGHTS for a job for the ABC client which are different from the XYZ client, you can create directories called ABC and XYZ, each with their own, different values. If you then set the Parameter Files Path to ABC, it will use these PARAMS and WEIGHTS, but if there is no PERFORM file in the ABC directory, it will use the one in the Database sub directory of matchIT.

Current Import path

The Current Import path from the matchIT Options, File Locations screen allows you to specify a different directory for matchIT to look in for a Source File to open. When you Import a Source File or open a Source File in the Setup Wizard, matchIT updates this path to the directory that you select the new Source File from.

Current Main File path

The Current Main File path from the matchIT Options, File Locations screen allows you to specify a different directory for matchIT to look in for a Main File to open. When you open a

Main File from the menu or toolbar, or save a new Main File layout in the Setup Wizard, matchIT updates this path to the directory that you select the Main File from.

Source File sub-folder

The Source File sub-folder from the matchIT Options, File Locations screen allows you to specify a sub-folder of a "job folder" that you use for the data that you run matchIT on. For example, if you use a standard name for a job folder of JOB12345 with a standard sub-folder name for source files of SOURCE, enter the name SOURCE in this option.

Main File sub-folder

When you save a new Main File layout in the Setup Wizard, matchIT will default to save the file in a sub-folder of this name, either within the directory that you select the Source File from, or within the directory above that if the Source File is in a sub-folder matching the name in the Source File sub-folder described above. For example, if you use a standard name for a job folder of JOB12345 with a standard sub-folder name for matchIT files of WORK, enter the name WORK in this option.

Temporary files path

The Temporary Files path from the matchIT Options, File Locations screen allows you to specify the directory that matchIT uses to store temporary files. You can specify this path when you install matchIT, but change it at any time using this option – you must restart matchIT for it to take effect. You should select a directory on a local hard disk – if you have more than one local hard disk, we recommend that you store temporary files on a different physical hard disk from that on which matchIT is installed. matchIT creates a lot of temporary files, which are usually cleaned up when you close matchIT, but if it crashes, they will not be deleted.

DAT files path

This option specifies where the primary matchIT data files are located. This includes, but is not limited to, the following types of information.

- Single word lookup (Names & Words table)
- Two word lookup (Names & Words table)
- Business Noise Words (Names & Words table)
- etc.

Other Options

The other parameter screens are described in their relevant sections:

- "Input Options" on page 33
- "Matching Setup" on page 97
- "US Address and Zip code Validation" on page 49

Multiple Parameter Sets

Save/Restore Setup

If at any time you want to revert to a standard configuration for the Matching Weights, Matching Matrices, or Option settings, you can do this from the Save/Restore Setup option from the Jobs/Setup menu. You can choose to individually restore any combination of the aforementioned settings by setting the check boxes on this screen. There are six types of matching for which relevant settings can be restored: Individual, Business, Household, Family, Company only, and Custom. Use the drop-down boxes labeled "for … configuration" and "for … nationality" to set the level of matching and the nationality of data. Finally, click on the 'Restore setup' button to restore the chosen settings.

If you are changing the nationality of data that you are processing, restore standard parameters for that nationality **before** you go into the Setup Wizard. If you are processing mixed country non-UK data, select **Worldwide** nationality. If you are processing both UK and non-UK data, select **Mixed incl. UK** nationality. If you are processing data all for one country, select that country if it is available in the list, otherwise select **Worldwide**.

Save User Parameters

If you want to save any of the settings that you have specifically set e.g. because they work well on a certain type of data, you can also do this from the Save/Restore Setup option from the Jobs/Setup menu. Set the check boxes according to which type of setting you want to save, then set the drop-down box under the check boxes to say "Custom"; the 'Create/update setup' button will then become enabled. The custom parameter set is indicated by a file extension, so type a file extension (which can be more than three characters) in the text box labeled "with File Extension".

Certain file extensions are reserved for the five standard matching setups; these are "IND", for personal matching, "BIZ", for business matching, "HH", for household matching, and "FAM", for family matching, "CO", for company only matching.

Restore User Parameters

To restore a set of previously saved user parameters, choose the Save/Restore Setup option from the Jobs/Setup menu, and set the drop-down box under the set of check boxes to say "Custom". Then type in the relevant file extension in the text box labeled "File Extension", click on the 'Restore setup' button and the chosen settings will be restored. If you created that parameter set for a different nationality, you must restore the standard parameters for that nationality before you restore the custom set.

Data Archival of Jobs

The 'Data Archival of Jobs' functionality is designed to archive all customizable files and parameters within a matchIT installation, to a location chosen by the user. These files can then be restored at a later point, in order to revert back to the exact matchIT environment, as it was at the time when the files were archived.

Upon restoring from an archive, the user will in theory be able to repeat a job that was carried out in matchIT directly before the archive was created, and get exactly the same results.

An archive might also be created, in order to save files that are specific to a company's data processing standards. For example, the users company may always need to have certain exclusion words included in the Names and Words table or use company specific report branding. Therefore, an archive can be created on a shared drive, and each time a user installs matchIT, they can restore from the shared archive before they begin using the new installation.

The files that are saved/restored by the data archival functionality can be customized by the user.

Creating an Archive

Before creating a new archive, the folder in which the archive is to be created must already exist.

To create a new archive of files, select 'Save/Restore Setup' from matchIT's Job/Setup menu. The 'Save/Restore Setup' dialog will appear. After selecting the "Advanced" button the dialog below will be displayed.
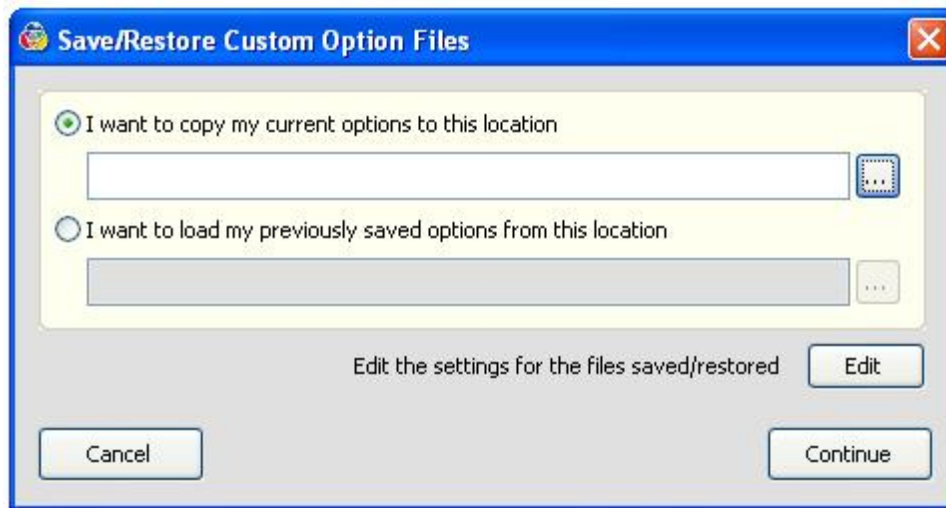
Make sure that the 'I want to copy my current options to this location' radio button is selected, and click on the corresponding ▣. A 'Select Directory' dialog will now appear. From here, navigate to the folder in which the archive is to be created, select that folder, and then click on the 'Select' button.

The 'Save/Restore Custom Option Files' dialog will now be displayed again, but with the archive folder selected. To create the archive, click on 'Continue'.

While creating the archive, matchIT will display information in the main Feedback window on which files are being archiving and the general progress. The overall process should not take more than a few seconds.

Restoring from an Archive

To restore from an archive, navigate to the 'Save/Restore Custom Option Files' dialog, but this time select the 'I want to load my previously saved options from this location' dialog. Click on the corresponding ▣, and this time select the folder containing a previously created archive. Upon being returned to the 'Save/Restore Custom Option Files' dialog, click on continue to begin the restore process.

Before the restore process starts, a prompt will appear asking for confirmation to proceed with overwriting the matchIT work file that was in use at the time of taking the archive, with an empty structure of that file. Choose 'OK' to proceed or 'Cancel' to prevent this file being overwritten.

Similar to creating an archive, the feedback window will display information on what is being restored, and general progress of the restore. The restore process should not take more than a few seconds to complete.

Editing the Files that are Saved/Restored

To edit the files that matchIT Saves/Restores, click on the 'Edit' button from within the 'Save/Restore Custom Option Files' dialog. A table (Restorelist.dbf) will be displayed. Each row in this table represents a different combination of folder to save from/files to save/folder to archive those files to.

The names of the folders to archive from/to are stored within the 'Location' field. The files within that folder that are to be backed up are specified in the 'Filename' field, and the folder in which they are to be archived to is specified in the 'Storeloc' field.

Wildcards can be used to specify more than one file, files with the same name but different extensions, or file with different names but the same extension. By default, the first record in the table represents matchIT's 'DAT' folder. In this, the 'Location' field contains '…\DAT'. The 'Storeloc' field contains 'DAT files' and the 'Filename' field contains '*.dat'. Therefore, by default, all '.dat' files from matchIT's 'DAT' folder will be restored to a new folder called 'DAT files' in the archive folder specified by the user. All '.dat' files in this folder will be restored back to matchIT's 'DAT' folder during the 'Restore' process.

To remove any records from this table, click on the small deletion cell. To indicate that a record has been deleted, this cell will turn back. The image below demonstrates how it would look if the second and fifth records were deleted.
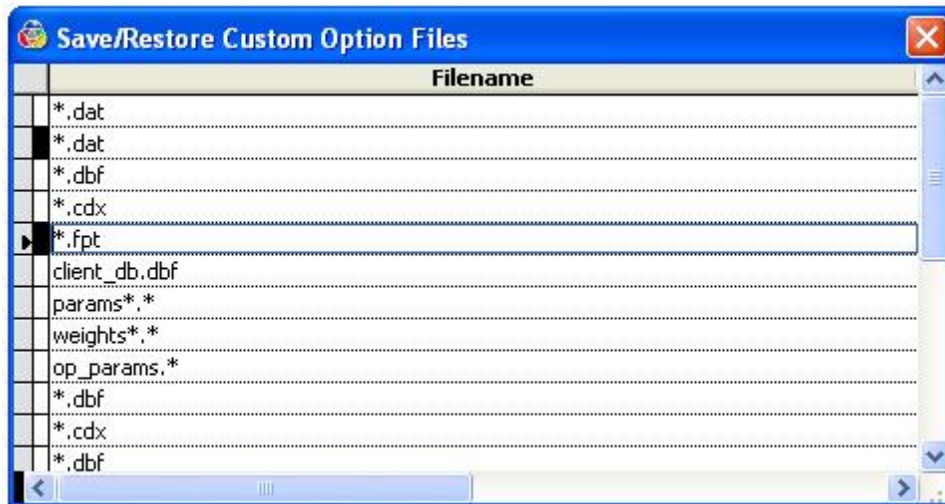
To add a new record, select 'Append New Record' from the 'Table' menu, whilst browsing this table.

Running Backup or Restore from a Job Script

Files can also be archived or restored from within a job script by mapping the program "backup_or_restore.prg" from the script. This program is supplied with matchIT and can found in matchIT's Progs sub-folder.

The program will need be mapped from a 'Source File' cell whilst browsing the script in the Editor screen. After mapping the program, the first two lines of active code will need to be edited by double-clicking on the corresponding Source File cell and selecting 'Edit Source File'. These lines

determine whether to run the 'Archive' process or the 'Backup' process, and where to archive to or backup from. The lines are as follows:

ProcedureToUse = "BACKUP"

BackupOrRestoreDirectory = "FULL PATH OF BACKUP DIRECTORY"

The colored text above is the code that will need editing.

After mapping the program from a job script and editing the program accordingly, the script can be run. The same feedback will indicate the success of this process.

Note: job script tables (Job and Job_Set files) cannot be archived or restored whilst calling this process from a job script.

Troubleshooting

- Feedback indicates that files could not be copied, during the 'Save' process:

    If any changes have been made to Restorelist.dbf, make sure that the changes point to valid files.

    Make sure that none of the files that Restorelist.dbf has mapped are opened or locked by any other applications or users.

    Make sure that the location in which the backup is being create can be written to, and that none of the sub-folders or files that matchIT may be trying to overwrite are opened or locked by any other applications or users.

- Feedback indicates that files could not be restored, during the 'Restore' process.

    If any changes have been made to Restorelist.dbf since the backup was created, make sure that the changes point to valid files.

    Make sure that the backup folder and its sub-folders that the backup was originally created in are still valid and have not been renamed, moved of deleted.

    If the backup folder is on a network, make sure that the location can be accessed freely and that the connection is OK. If in doubt, try copying files from that location manually.

Setup Wizard

Launching the Wizard

You can start any of the Setup Wizard's from matchIT's Wizard Pane, or you can access the Setup Wizard from the Import menu. The Getting Started Guide provides a step by step example for using the Setup Wizard.

Importing Data

Assuming that you have a file of names and addresses, you need to specify its Input Format and the Field Layout. If the file is on CD or some other removable media, it may be best to copy it to your hard disk first, making it ready for input.

The Setup Wizard (on page 20) provides the simplest method of importing data into matchIT, from a variety of file types. The Wizard will read DBF, Access, Excel and ASCII text files, as well as an ODBC Connection file. Text files can be one of four types: comma delimited (commonly known as CSV), tab delimited, fixed width or fixed width without Carriage Return Line Feed (CRLF) characters – for this last type, you must specify the record length.

The use of the Setup Wizard is described in the Getting Started Guide, which is available as a separate booklet or in PDF format from the matchIT Help menu.

Effective Quality Assurance is necessary to detect any problems with less common or inconsistent input formats.

If files are inconsistent in their input format, they will not be imported properly. Please note that the Intelligent Setup Wizard only examines and displays the first few records in a file when trying to determine the type of data and its width – make sure that (for delimited files) fields are not being truncated. It is best to err on the side of caution by allowing too much width. You can check or change maximum field widths after Import by using the Database Utilities option Change Fields to Fit (on page 164).

Before Import

After clicking "continue", the next screen in the Setup Wizard asks what needs to be done with the data next. The exact options available will depend upon the version of matchIT being used and the type of data imported.
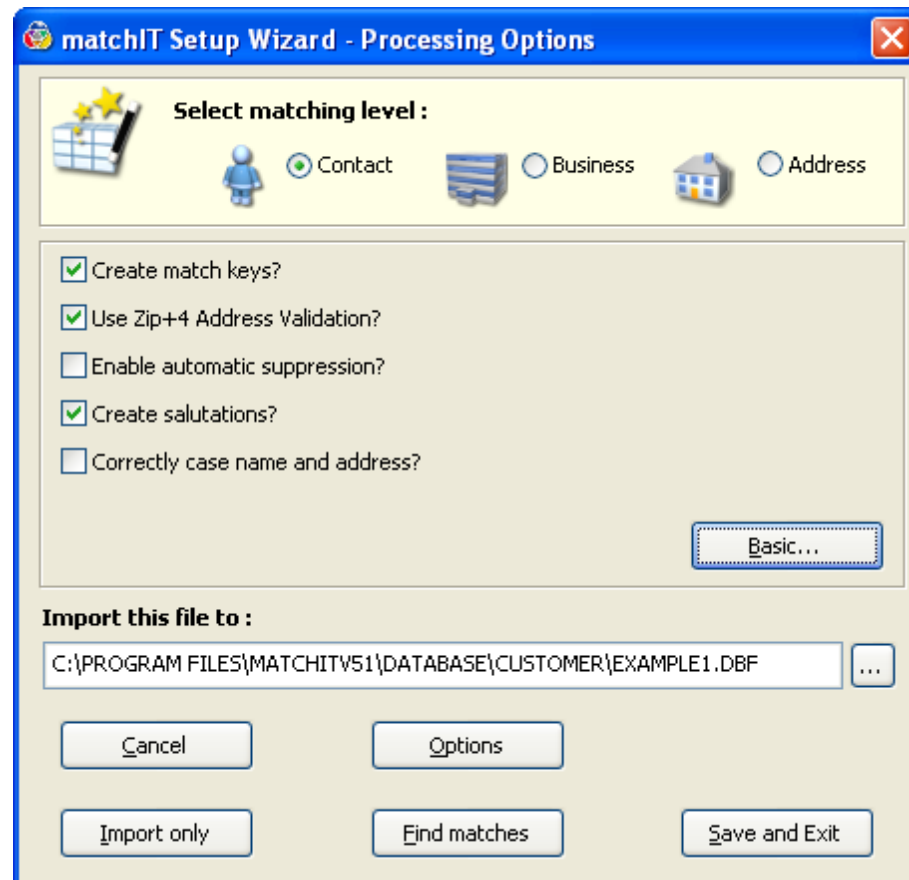
- "Create Match Keys" must be checked to find duplicates within a file, or find matching records in another file e.g. if this file is a suppression or "stop" file.

- "Use Zip+4 Address Validation" can be checked to verify and correct US zip codes and address (available only with addressIT).

- "Enable Automatic Suppression" can be checked to verify and remove Suppression records (available only with **suppressIT).**

- **"**Create Salutations " must be checked to derive Salutations (e.g. Dear Mr Brown)- these can only be derived if Match Keys are created. The Default Salutation (see Default Salutation) is used if matchIT cannot work out the correct salutation. You can change the default using the Options button.

- "Correctly case name and address" (available only in matchIT **Pro** or with the **Salutations and Casing** module) should be checked if:

  - the data is all upper case

  - the casing is poor or inconsistent

- Matching records can be found to:

- Individual, Family or Household (Address) level for personal data

- Contact, Business or Address level for business data

- You must select the matching level by using the buttons at the top of the window.

- The More…/Basic… button toggles between advanced import options and basic import options.

What happens after the Setup Wizard is determined by the options set and then which action button is selected.

Import Only

Select this to create a Main File and import the input data file into it, without finding matches. Use this if:

- you want to change any of the matching rules before Find Matches

- this is a suppression file which you want to compare against another file but don't want to Find Matches within.

Find Matches

Select this to create a Main File, import the data and then find matching records using default matching rules. Don't select this option if you are setting up this Main File to use in the Multiple File Wizard later, because the Multiple File Wizard performs the matching.

Save and Exit

Select this to create a Main File layout with the required fields (including matchIT's fields chosen through the checkboxes), but empty of records. Use this if you want to change any Input or Operational options before importing the file, or if you are setting up Main File layouts to use in the Multiple File Wizard (on page 125) and want to leave the time-consuming process of Import until later.

Cancel

Select this to quit the Setup Wizard and lose the information entered about the file setup.

End of Import

Once the data has been imported, matchIT displays a dialog to allow you to choose from several options:
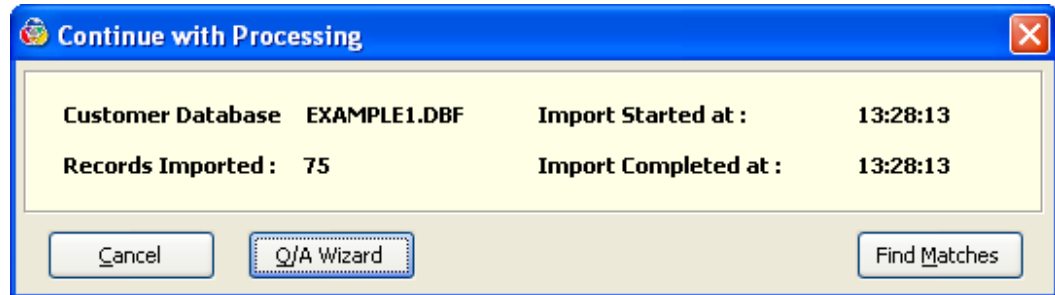
- "Q/A Wizard" to view the imported records, data summaries, records by categories, and rejected records.

- "Find Matches" to display the Matching Key and Range screen.

- "Cancel" to return to the main menu.

Importing Data

Importing without using the Setup Wizard

If the file is in the same input format and field layout as a file that you have previously loaded into matchIT, you don't have to go through the Setup Wizard again. The field layout is stored in the Main File that you used previously and the Input Options specify the Input Format and what processing you want to do on Import.

To import a file into matchIT without using the Setup Wizard, select Import Records from the Import menu. Firstly, select the Main File that you used previously, or create a copy of this file's layout by using the option. matchIT then prompts you to change the Input Options if necessary (via the Change Basic Parameters button) e.g. if the last file you Imported was Comma Delimited but this file is Fixed Width:
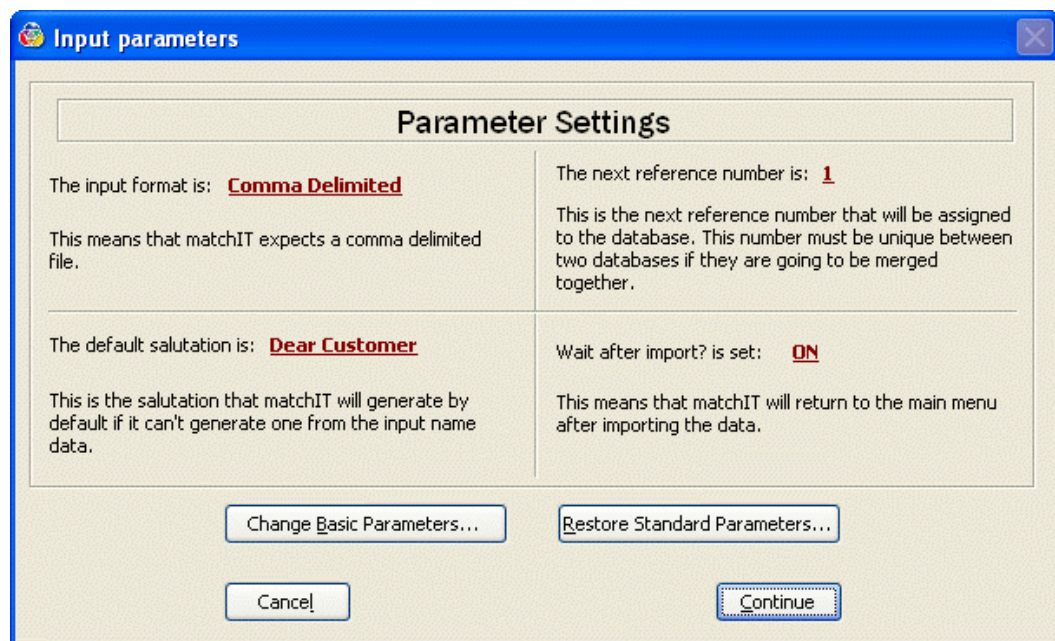
If you want to dedupe the file at a different matching level than the last file you deduped e.g. individual last time but family level this time, select Restore Standard Parameters and select the appropriate matching level. This restores not only the matching level (in Matching Options) but also the matching weights (and the settings for a different nationality).

Next, select the input file (the source file) that you want to load into matchIT.

If there are records in the Main File from a previous run, matchIT wipes the file clean (also known as "zapping" it) before it loads in the new data.

Contacts and Salutations

This section is only applicable if you have matchIT **Pro**, or the Salutations & Casing module.

matchIT **Pro** has the facility to generate contacts (Prefix Initials Surname e.g. Mr J R Smith) and salutations (e.g. Dear Mr Smith) for every record in a Main File. To derive salutations, check the "Create Salutations?" box at the end of the Setup Wizard (on page 20).

If your data is originally in "prefix, forename, surname" format, matchIT will use these fields to generate the contact. Otherwise it will look for a prefix in the "addressee" field or use the forename to assign Mr or Ms if there is no prefix present – it can only do this when the forename is in matchIT's Names and Words table as male or female. Prefixes are also defined by entries in matchIT's Names and Words table, along with the appropriate salutation rules.

A default salutation is used when matchIT cannot work out an appropriate one from the customer's name. This default can be changed via the Default Salutation (on page 35) option under the Jobs/Setup menu, Options, Input tab.
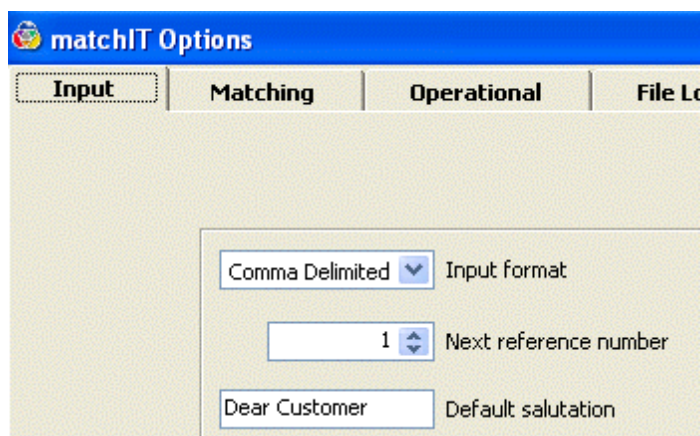
Names and Words

The Name and Words facility (from the Jobs/Setup menu) is used for changing or adding entries to the NAMES lookup table. You can choose to add an entry, remove an existing one, or search the table for a particular entry. For adding an entry, there is another screen where the word to add can be entered, along with the matching equivalent (which matchIT uses when matching), the type (chosen from a drop-down list), the sex and salutation (if appropriate), and the proper case value (which matchIT will use when casing).

When searching for an existing word, there are two choices; 'Search' and 'Search Rest'. 'Search' will find the first occurrence of the entered word, while 'Search Rest' finds the next occurrence from the current position in the table, this enables multiple entries to be found. The word can be searched for as a specific type, or as any type.

The NAMES lookup table is actually held in a fixed-width text file in the DAT sub-directory; the layout of this file is as follows:

Heading	Width	Description
TYPE	1	Type of entry – see below
EQUIVALENT	25	Matching equivalent of the entry (e.g. 'Tony' has a matching equivalent of 'Anthony', enabling these two names to be matched)
NAME	10	The word which is actually looked up
SEX	1	Indicates the sex of the forename or prefix
SALUTATION	2	Indicates the type of salutation to be generated for a particular prefix
PROPER CASE	30	Proper case value for the entry
SWITCH	1	Indicates whether this entry is the first part of a double word phrase

The different "Types" that can be entered in the table are as follows:

Type	Description
'A'	Address Word, such as "Rd" or "Street"
'B'	Business word, such as "Ltd" or "Printers"
'C'	UK county, such as "Kent" or "Glos"
'E'	Exclusion word, such as "Deceased" or "Moved"
'F'	Female forename (note the sex has to be set for these entries too)
'J'	Job title word, such as "Manager"
'M'	Male forename (note the sex has to be set for these entries too)
'N'	Noise word (i.e. ignored when generating keys or address matching), such as "The" or "House"
'O'	Overseas i.e. foreign country
'L'	Local country, such as "UK" or "Scotland"; this enables local countries to be identified as countries, without the record being marked as foreign
'P'	Prefix, such as "Mr" or "Captain" (note the sex has to be set for these entries too, also the SALUTATION TYPE)
'Q'	Qualification word, such as "PhD" or "ARICS"; these entries typically always need a

Type	Description
	proper case entry as casing of qualifications can be unusual
'S'	Special casing word, i.e. a word that is cased unusually but doesn't fall into any of the above categories, such as "PhotoMe"
'T'	State or province, such as "Pennsylvania" or "PA"
'U'	Unknown word; this is for the first words of a two word phrase, which, on their own, have no special meaning, such as the "Hong" in "Hong Kong"

Each prefix entry must have a salutation type associated with it. The following list shows the salutation types, along with an example of the type of salutation that will be generated:

Type	Rule	Example
S	Dear Prefix Surname	Dear Mr Smith
C	Dear Prefix Surname	Dear Mr Smith
FS	Dear Prefix Forename Surname	Dear Mr John Smith
FF	Dear Forename	Dear John
F	Dear Prefix Forename	Dear Sir John
B	Dear Prefix	Dear Sir
T	Prefix	My Lord

In addition to these, the following salutation types will effect the way in which matchIT will generate a contact as well as a salutation.

Type	Salutation Rule	Example
SZ (Suffix)	Dear (def) Prefix Surname	Dear Mr Smith
ZZ (Suffix)	Dear (def) Prefix Surname	Dear Mr Smith
TA	Prefix	Your Eminence
BA	Dear Prefix	Dear Madam
SA	Dear Prefix Surname	Dear Family Smith

*'SA' is the same as type 'S', but if a full forename is found, the full forename is ALWAYS included in the generated contact, regardless of whether or not the 'Use full name in contact?' (on page 35) option is ticked.

Type	Contact Rule	Example
SZ	(def) Prefix Forename Surnames Suffix	Mr John Smith Jr
ZZ	(def) Prefix Forename Surname	Mr John Smith
T A	Prefix Forename Surname	Your Eminence John Smith
BA	(EQUIV) Forename Surname	Countess Jane Smith
SA	Prefix Forename Surname	Family John Smith

Additionally, each prefix, male forename and female forename must have a sex associated with it, taking a value of either 'M' (Male), 'F' (Female), or 'E' (Either).

Import Records

matchIT first reads the records from the input file into the selected Main File. If they are already in a matchIT standard DBF file (see "Main File Layout" on page 30), then you can select Generate Keys from the Import menu instead of Import. Both options perform processing as follows (assuming that the relevant fields are defined in the Main File Layout and that you have the appropriate modules of the matchIT suite):

- generate salutations and phonetic keys from the various name fields

- splits freeform personal names into component parts or combines them

- generate phonetic or non-phonetic address keys

The processing described above is implicitly specified by field names in the Main File Layout – the Setup Wizard adds the fields when it creates the file, depending on the options chosen. The processing described below is controlled by the Options chosen at the end of the Setup Wizard or on the Options screen (in the Jobs/Setup menu):

- correct and validate zip codes and addresses

- convert names, address lines and (if present) job titles to "proper case"

- relocate data items to appropriate fields – this can include zip codes, company names and job titles.

Information about the progress of these operations appears according to the Progress Interval specified – select the Operational Tab from the Options screen, Job/Setup menu, to modify the Progress Interval. After every Progress Interval, the display is updated. If the Main File has more records than "Max database size to display stats on info screen", a Browse view of the data being processed is displayed on the right hand side of the screen. The top half of this display shows fields from the input data, in the input order. The bottom half shows matchIT's generated fields, in an order specified by a parameter file (IMPORTFIELDS.DBF).

You can pause at any stage by pressing the Escape key i.e. if you want to devote processing power to another task for a while, or if you want after changing or.to change the Progress Interval or abort the Import leaving the progress interval, press Esc or F3 and then select 'Yes' to abort processing or 'No' to continue.

If you abort and later want to continue from where you left off, select the Generate Keys option from the Import menu. You will be prompted with the record number that follows the last record

to be fully processed before the interruption. Check "Start at beginning" to generate keys from the beginning of the file; otherwise, confirm the number of the last record for which keys have been generated, or enter a different value.

During Import, matchIT writes a copy of the progress information appearing in the Feedback Window to the PERFORM.TXT file in the same directory as the Main File. See "Performance Data" (on page 9) for more details.

After carrying out these operations, if the Wait for User after Import option is ON, matchIT will ask if you want to use the Q/A Wizard View (to View Data or View the Data Summary)or Find Matches.

If the "Wait for User after Import" option is not checked, after Import has finished matchIT will automatically Find Matches using the default Match Keys.

The Data Summary will provide information on the newly imported data.

Data Summary

This is an option under the matchIT Import menu or via the Quality Assurance option under the Output menu – from here, you can choose to preview the report, print it or output it to a file. This option is described in more depth under Quality Assurance Introduction (on page 138).

View Data

This option is available as a hyperlink at the top of matchIT's main window, or within the **View** menu as **View Data in Main File**. View Data allows you to browse and update the selected Main File in any order. It is similar to the View menu, **View Table** option, except that you don't have to specify the table to browse and you can browse the Main File in order of any field within it.

First, an options screen will be presented:
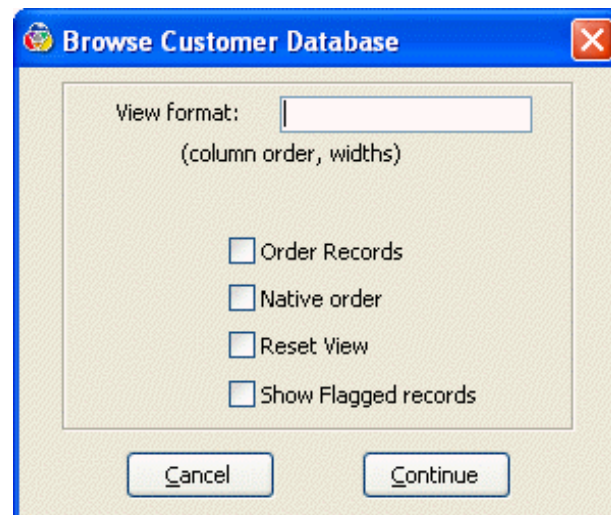
- Select the **View Format** you want to use, or enter the name of a new format under which to save the view that you create next. You can manipulate the display as normal.

- **Order Records** allows you to choose a field to order the records on (select a field from the list presented next). Each time you select to browse in order of a field for which an index was **not** created at the end of Import (which usually includes the default Match

Keys and UNIQUE_REF), matchIT will create an index – this will take quite some time on a large file.

- **Native View** means the fields will be displayed in the order they are in the Main File – the browse view may have been previously changed by changing the files sort order or column order.

- **Reset View** means the column order and width will be reset in the browse view. – the browse view may have been previously changed by changing column order and widths.

- **Show Flagged records** will include records that were flagged from the Main File (usually by Flag Matches) in the browse view. These will have a black marker in the leftmost column.

Main File Layout

Summary

The best way to configure a Main File with all the fields required for processing your data file is to use the Setup Wizard (on page 20). If, for some reason this is not possible or not an appropriate option, this section explains the fields your Main File must have so that matchIT can do what is required.

There are basically two kinds of fields that the system uses. First, there are **Input Fields**. These are the fields that you, the user, will supply for matchIT's processing. If they are name and address fields, you must give them standard matchIT field names, because otherwise matchIT will not recognize that it must apply appropriate processing to these fields. For example, it must know if a field is a freeform field containing a person's name, or if it contains just the surname, or if it is an address line etc. Although matchIT can search through name fields for job titles, through address lines for names or zip codes etc., it will work much more effectively if you can label your data with appropriate matchIT standard field names.

Additionally, matchIT uses **Generated Fields**. These are fields that probably weren't in your input data, but matchIT will generate them from your data. They are critical to the main processing steps (deduplication, etc). It is normal to have the input fields at the start of your Main File layout, in one block, and to have the generated fields in another block at the end of the layout. We recommend that you keep to this unless you have a special reason not to, as this allows matchIT to work out which fields are Input Fields and which are Generated.

However, as shown in matchIT Field Names and Usage, some fields can be Input in one file but Generated in another. Also, some Input fields can be replaced by Generated fields.

matchIT Field Names and Usage

The table in Appendix A - Field Names & Usage (on page 174) lists field names and usage in alphabetical order.

Quality Assurance Field

matchIT requires a FLAG field to prepare the Data Summary and for View Records by Category to allow you to examine certain categories of a record. Entries in the Flag field are explained in Appendix B on page 189.

matchIT Field names by Type

These are the main data fields that matchIT recognizes specifically for import or matching:

Personal Name

- ADDRESSEE, CONTACT, FORENAMES, FORENAMES2, INITIALS, INITIALS2, NAME, NAME1, NAME2, NAME3, NAME2FOUND, NAME_KEY, PREFIX, PREFIX2, QUAL, SALUTATION, SEX, SUFFIX, SURNAME, SURNAME2

Company Name

- COMPANY, COY_KEY, COY_NAME, NAME1, NAME2, NAME3

Job

- DEPARTMENT, JOB_TITLE

Address

- ADDRESS1 to 9, ADD_KEY, COUNTY, COUNTRY, DPS, POSTCODE, POST_IN, POST_OUT, PREMISE, STATE, TOWN, ZIP

Telephone

- FAX, REST_PHONE, STD_CODE, TELEPHONE

Campaign

- MAILING_ID, MAIL_DATE

Match Set Grouping

- MATCH_REF, SET_DUPS, OVERLAPREF, OVERLAPCNT, MATCHSCORE

Deletion Priority

- DEL_PRI

Quality Assurance

- FLAG

Unique Reference

- UNIQUE_REF

Using Your Own Field Names

You may use your own field names within matchIT. If you do, matchIT will not recognize them for special processing (e.g. if you name a field as CUSTOMER instead of ADDRESSEE, matchIT will not apply its personal name processing to that field). This is because matchIT uses the field name as the main indicator to tell it what kind of data is likely to be in that field.

Although matchIT won't apply any special processing to your own field names, it will apply its standard string handling routines. This means that you can use any field you like as either a matching key or for matching weights. We don't recommend doing this until you are confident that you completely understand how matchIT works, as using a non standard field can have a drastic effect on your matching results.

To allow you to use your own field names safely, we have set down a few guidelines on what to call your fields. If you follow these rules, you should have no problems using your own fields in matchIT:

- Each field name in a Main File should be unique. If you have fields that do a similar thing, number them – for example, "CUSTOMER1", "CUSTOMER2" etc.

- Whenever possible, select the field name "OTHER" from the suggested list in the Setup Wizard for fields that you know matchIT won't need to know about. We appreciate that in many situations, these names would not be sufficiently meaningful for processing, so in these cases, follow the naming rules below.

- Field names must not be greater than 10 (ten) characters long

- Field names must ONLY contain the letters A-Z or the numbers 0-9, or the underscore character ("_"). Field names must NOT contain spaces or any other kind of punctuation.

- Field names must start with a letter (A-Z) NOT a number or an underscore.

- Field names are not case sensitive, so 'CUSTNAME' is the same as 'CustName' or 'custname'

- Field names must be AT LEAST three characters long.

- Although you should follow these rules for your field NAMES, you can use different TYPES of fields (character, numeric, logical, date etc.). However, matchIT can only match on character fields.

- Do not use any field names specified in the reserved list, as use of these names may conflict with matchIT's operation.

Words Not To Use As Field Names

The Online Help contains a list of "reserved names" which you should not use as field names.

Input Options

Basic Options

Input Format

Input Format is a drop down list from which you can select as follows:

- **Comma Delimited or COMMA** (commonly known as CSV): Use this setting for files that are either Comma or Comma and Quote delimited. This format refers to an ASCII text file in which records can be of variable length, ending with a carriage return and line feed. Commas separate the fields, and additionally double quotation marks can delimit character fields. For example:

    - "MR G. SMITH",456620,"123 CHURCH ROAD"

    - If the input data does not have the double quotation marks for character fields, this causes a problem if there are commas within a field e.g. Flat 103, 10 High Street intended as one address line. Therefore, this format cannot be guaranteed to be imported 100% successfully.

- **Fixed Width or SDF** (System Data Format): refers to an ASCII text file in which the fields and records all have fixed lengths, there being no field delimiters, and a carriage return and line feed as the record delimiter. This is a common choice for data coming from IBM and plug-compatible mainframes.

- **Tab Delimited or TAB**: Tab format is for an ASCII text file in which records can be variable length, ending with a carriage return and line feed. Fields are separated by the

TAB character and can be additionally delimited by double quotation marks. This format is preferable to Comma Delimited, as there are unlikely to be problems if the fields are not surrounded by double quotation marks.

- **PIPE**: Pipe format is identical to tab delimited, except that the fields are delimited by the pipe ("|") character. This will not work unless the fields are also surrounded by double quotation marks.

- **DBF Table or DBF**: Database File format refers to a dBase type file i.e. one conforming to the dBase / FoxPro DBF file standards. matchIT itself uses this standard for all its tables. If you select this format, you will then be prompted to indicate whether the DBF file containing the data conforms to matchIT's field naming standards (see "Main File Layout" on page 30).

  - If you reply "No", then you must create a skeleton DBF which does conform to matchIT's field naming standards, into which matchIT will copy the data from your DBF. During Import the data from the input DBF file goes first to an intermediate temporary DBF file, to enable matchIT to transfer it to a file conforming to matchIT naming standards.

  - If you reply "Yes", then matchIT will change the Input Format from DBF to EXT to tell it just to extract or re-extract Match Keys, salutations etc. without importing the data into the matchIT Main File. You can use EXT format in a Job Script to extract or re-extract Match Keys in a Main File. If you are importing/extracting through the Import menu, you can use Generate keys to do this without changing the input format.

- **Access Table or MDB**: Access format; you will be prompted for which of the available tables in the database you want to import.

- **Excel Worksheet**: Excel v5.0, v6.0 or v7.0 format; note that if there are multiple worksheets, matchIT will only be able to read the first sheet. Use ODBC for multiple worksheets with the data to be imported being a defined name in Excel.

- **Regenerate Keys or EXT**: DBF format conforming to matchIT's field naming standards (see DBF input above).

- **DIF**: Data Interchange Format used by VisiCalc.

- **FW2**: Framework II files.

- **MOD**: Microsoft's Multiplan version 4.01.

- **PDOX**: Borland's Paradox version 3.5 or 4.0 database files.

- **RPD**: RapidFile version 1.2.

- **SYLK**: SYmbolic LinK interchange format used by Microsoft's Multiplan.

- **WK1**: Lotus 1-2-3 spreadsheet revision 2.x.

- **WK3**: Lotus 1-2-3 spreadsheet revision 3.x.

- **WKS**: Lotus 1-2-3 spreadsheet revision 1.A.

- **WR1**: Lotus Symphony spreadsheet versions 1.1 or 1.2.

- **WRK**: Lotus Symphony spreadsheet version 1.0.

- **XLS**: Microsoft Excel version 2.0, 3.0 or 4.0.

- **No CRLF**: Select this option for an ASCII text file in which records are of fixed length, not ending with a carriage return and line feed, or other specific separator. If you select this format, you will then be prompted to enter the number of characters in each record. During the Import operation, the data goes first to an intermediate flat file with a CRLF

terminator for each record. The name of this file is SDFnnn.TXT (where nnn is the length of the record), and matchIT will delete this file when exiting. If the file contains ASCII null characters instead of spaces, you should reply "Yes" when asked if you want to convert nulls to spaces.

- **ODBC**: Select this option for a remote data source. You will be prompted for an ODBC Connection file (on page 40). This file must be created prior to import.

Next Reference Number

The Next Reference Number parameter must be numeric. This will be used for a matchIT generated unique reference field called UNIQUE_REF. Enter in this field the number at which you wish matchIT to start sequential allocation of reference numbers, for example, 1. Enter zero if you want a prompt, asking you for a start number, just before Importing commences. If you are Importing two such files prior to merging them, this facility will allow you to make sure the reference numbers in the merged Main File don't overlap.

Default Salutation

This parameter determines the default salutation, either where matchIT can't determine one (for example, C Smith or Chris Smith, which could be either Mr or Ms), or where the Prefix imported isn't sex-specific (Dr, for example). If you include the word 'Dear' in the default salutation field on the Basic Parameters screen (e.g. actually key in "Dear Customer" and not just "Customer", so the screen shows Dear as a literal followed by "Dear Customer" in the data entry box), then all the salutations derived by matchIT will start with the word "Dear" unless the salutation for the type of title (or prefix) specifies "Title" only. For example, Mr J Smith will result in a salutation of "Dear Mr Smith" whereas The Bishop of Liverpool will result in a salutation of "My Lord".

Default Gender

The Default Gender parameter is the sex to assume when matchIT can't determine whether the name is male or female e.g. Chris Smith, C Smith. If you set this parameter to M or F, matchIT will assume it to be male or female accordingly, and develop a salutation using Mr or Ms as the prefix.

Extract Zip code

When processing address lines, matchIT can either extract (i.e. move) or copy "floating" ZIP codes from the address lines to a fixed field; set this parameter to "Extract" or "Copy" if required, or else set to "Leave". Note that a field labeled ZIP is required to put the data into – if there is no ZIP field identified during the Setup Wizard, it will automatically add the ZIP field to the main File Layout. If this field is not present, the setting of this parameter is ignored.

Use Full Name in Contact Field

Set this parameter "on" to include the full forename of any incoming name in the CONTACT field; just the initial will be used if the parameter is off. For example, if the parameter is on, and the incoming name is "John Smith", then the generated contact will be "Mr John Smith", if it is off, then the contact will be "Mr J Smith".

Omit Excluded Records

If matchIT finds Exclusion words such as "Deceased" in any field it is splitting up e.g. addressee, company name or address (if proper casing the address or using a phonetic address key), you can choose to automatically delete them during Import or Generate Keys. If you set this parameter "on", they are deleted, otherwise they are not. All excluded records are marked as such in the FLAG field (if present in the Main File Layout). Thus, if Omit Excluded Records is not switched on, you can choose to delete or mail these records after importing by selecting View Records by Category (on page 138).

Proper Case Incoming Data

If this parameter is set on, the Import step will convert the address lines in your records (labeled ADDRESS1, ADDRESS2... ADDRESSn) to their proper case. It will also convert ADDRESSEE, JOB_TITLE, DEPARTMENT, and COMPANY to the correct case. matchIT has a list of words it knows should be either all upper case, or all lower case, or a special mixture; these words can be found in the NAMES table, under the column PROP_CASE. This proper casing will handle punctuation, apostrophes and abbreviations.

matchIT's default rules for casing data are as follows: letters following an apostrophe are capitalized (e.g. "Mr O'Reilly"), as are letters following "Mc" or (subject to one of the Advanced Input Options) "Mac" at the start of a name (see "Mac Name Treatment" on page 39). Double-barreled names have a capital letter after the hyphen. If the name or other word has an entry in the PROP_CASE column in the Names table, it is cased as shown there e.g. BSc, helpIT, IBM, plc. If not in the Names table, words are all capitals if they contain no vowels, otherwise they are changed to initial capital followed by lower case letters.

To add new words that you would like matchIT to case differently from the rules above, or to change existing entries in the table, select Names and Words (on page 25) from the Jobs/Setup menu. The Equivalent field contains the word to be looked up, the Name field the matching equivalent (e.g. "Anthony" for "Tony"), and the Type field denotes what sort of word is being added. The Sex field is only used for forenames and prefixes, and the Salutation field is only used for prefixes. The Proper Case field is where any special casing of the word should be entered.

Nationality Options

Nationality of Data

As there are other settings (including match key defaults and weights) that need to be set for processing foreign data, you can't change Nationality via this screen: you must use the Save/Restore Setup (see Online help for more details) option under the Jobs/Setup menu to change the nationality settings.

This parameter controls whether the Setup Wizard labels the Postcode/Zip field as POSTCODE or ZIP in the Main File Layout, and (for Postcode) whether it triggers matchIT's postcode verification and extraction logic. If the country is UK, it will use the label POSTCODE and set the postcode logic on. Otherwise, it labels the field as ZIP and does nothing special or automatically with the contents of the field.

Extract Country

When processing address lines, matchIT can either extract (i.e. move) or copy "floating" countries from the address lines to a fixed field; set this parameter to "Extract" or "Copy" if required, or else set to "Leave". Note that a field labeled COUNTRY is required to put the data into – if this parameter is switched on before you use the Setup Wizard, it will automatically add the Country field to the main File Layout. If this field is not present, the setting of this parameter is ignored. Note also that the data in the field will be proper cased if that option is switched on – if you want it in upper case, you must amend the entries for the countries using the Names and Words option under the Jobs/Setup menu.

Extract Postcode

If the input file contains UK postcodes which are not in a fixed field, set this parameter to "Extract" to search through the address lines for a postcode and to move it to the postcode field. matchIT will only extract full postcodes, having an outward half which is valid according to the Mailsort tables. Otherwise, if the postcodes are in the postcode field, you can select "Leave" to leave postcodes as they are. You can also select "Copy" to copy the postcode from the address lines into a fixed field, but to leave it in the address lines as well.

Extract State/Province

When processing address lines, matchIT can either extract (i.e. move) or copy "floating" states and provinces from the address lines to a fixed field; set this parameter to "Extract" or "Copy" if required, or else set to "Leave". Note that a field labeled STATE is required to put the data into – if this parameter is switched on before you use the Setup Wizard, it will automatically add the State field to the main File Layout. If this field is not present, the setting of this parameter is ignored. Note also that the data in the field will be proper cased if that option is switched on.

Abbreviate States

Choose this option to have matchIT abbreviate States or Provinces when processing address lines e.g. to change Pennsylvania to PA.

Expect Building Number Before Apartment Number

This parameter is used when addresses processed by matchIT contain two numbers in the address line e.g. 12/24 High St. If this parameter is "on", then the first number is taken as a premise number and the second as a flat number, and vice versa if it is off. This is important when matching with the 'Must have premise match?' parameter on, or extracting or copying the premise. As a rule of thumb, set this parameter off if your file contains mainly UK addresses, otherwise turn it on.

Advanced Input Options

The Advanced Input Options are described in the Online Help.

ODBC Manager

matchIT can read and export data using standard ODBC connections. Supported ODBC drivers include the standard ones for SQL Server, mySQL, DB2, Oracle and Access.

When importing data, the user must create a single table or view containing the data which is to be read into matchIT – for Access, this can be a query. The user can select which fields from the table or view they would like to import. matchIT uses an ODBC Connection file to import the required information from the chosen data source, as well as export data back to the original source database. On output, the user can choose to create a new table, overwrite an existing table, or delete records from within a table and reassign orphaned records in a child table.

The following sections explain matchIT's ODBC functionality in detail, covering all of the major options and processes.

Creating an ODBC Connection file

You can create an ODBC Connection file by selecting the ODBC Connection Manager option from the File menu or by choosing the ODBC option in the matchIT Import Manager window, as shown below (to access the ODBC Connection Manager window - click "…" after choosing the ODBC radio button).
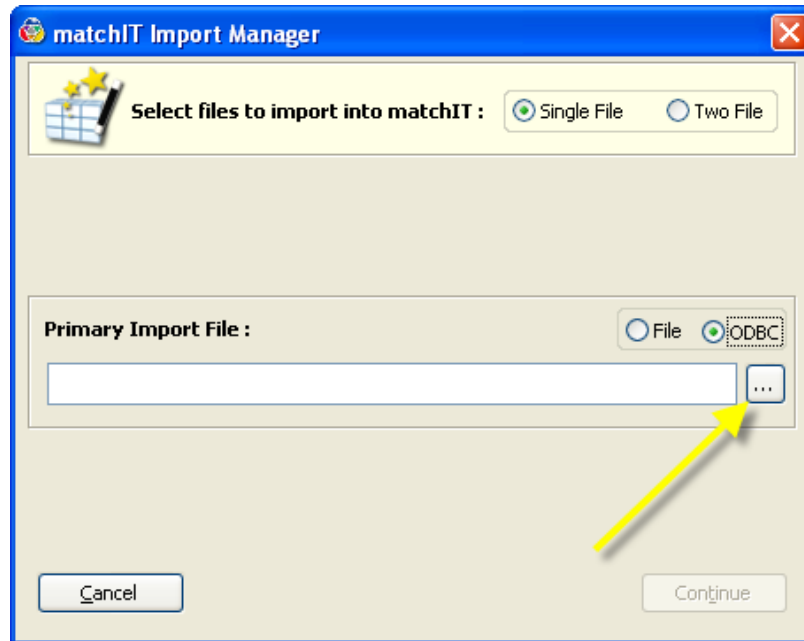
The ODBC Connection Manager window is then displayed as follows.
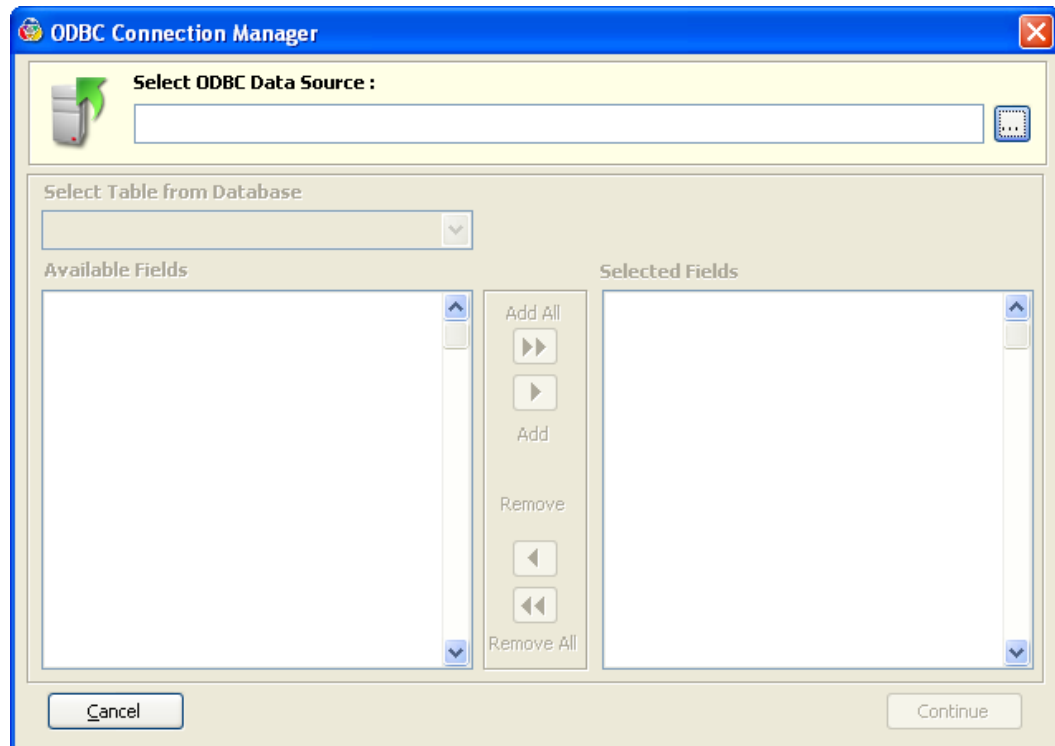
Choosing an ODBC Data Source

To choose an ODBC Data Source, click the ▣ button. If prompted, enter the Login ID and Password for the data source.

Selecting a Table

After choosing a data source, select a table or view from the drop-down list of available tables. This defines the data that will be imported.

Choosing Which Fields to Import

When importing a table or view, you can choose what fields you would like to import. This feature can help improve processing time by limiting the size of large files with numerous fields that are unrelated to the matching process, as well as reducing the setup time required in matchIT. You can select fields individually by double-clicking, select a block of fields by highlighting them and clicking "Add", or use the Add All button.
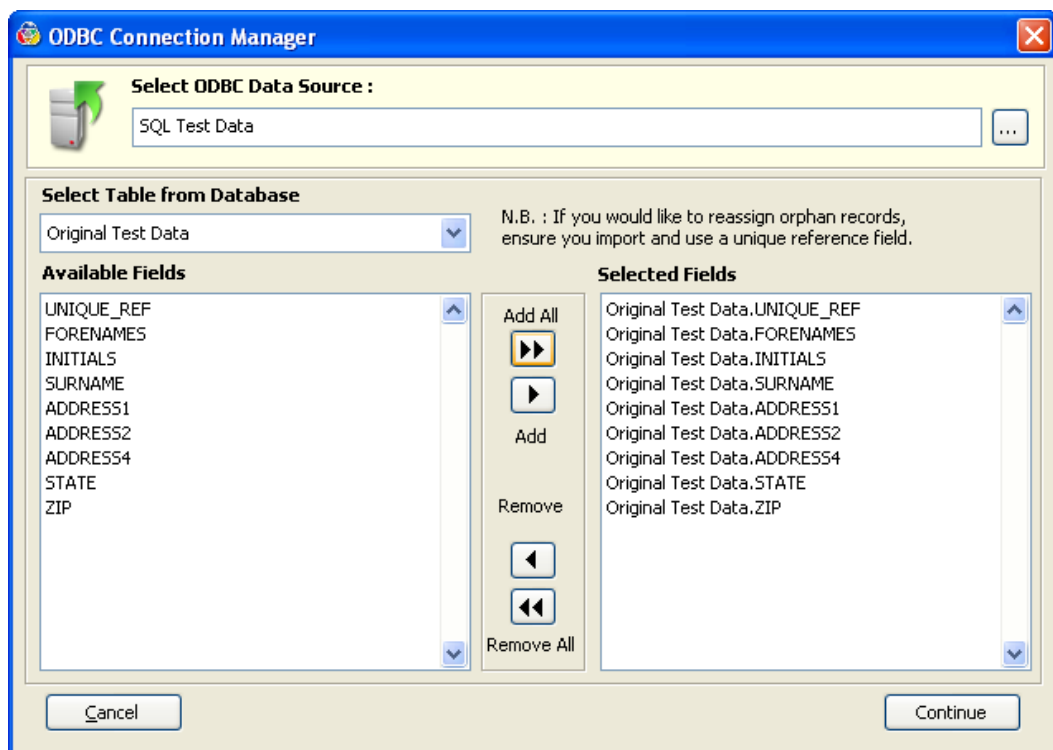
Importing via an ODBC Connection

Once you have created an ODBC Connection file, you can import the data selected via matchIT's Setup Wizard or by importing it into an existing matchIT Main File. Both methods are described below.

Using matchIT's Setup Wizard to Import ODBC Data

The matchIT Setup Wizard can be launched by selecting one of the file wizards or by selecting the Setup Wizard option located under the Import menu. Choosing the latter method will prompt you for the file that you wish to import, so you can only use this if you have previously created an ODBC Connection (*.ODB) file. However, starting the Setup Wizard via the Single File or Two File button will allow you to specify an ODBC Connection file or create a new one, as discussed in "Creating an ODBC Connection file".

Importing ODBC Data into an Existing Main File

To import ODBC data into an existing matchIT Main File, the DBF structure of the Main File must be consistent with the fields selected in the ODBC connection. If the field layout is not consistent between the DBF and the ODBC connection file, data will be misplaced causing problems during import or matching.

You can import records into an existing matchIT Main File by selecting the Import Records option from the Import menu. If the Main File is not already open, matchIT will prompt you for the file. Next, you will see the Input Parameters window, as described in .
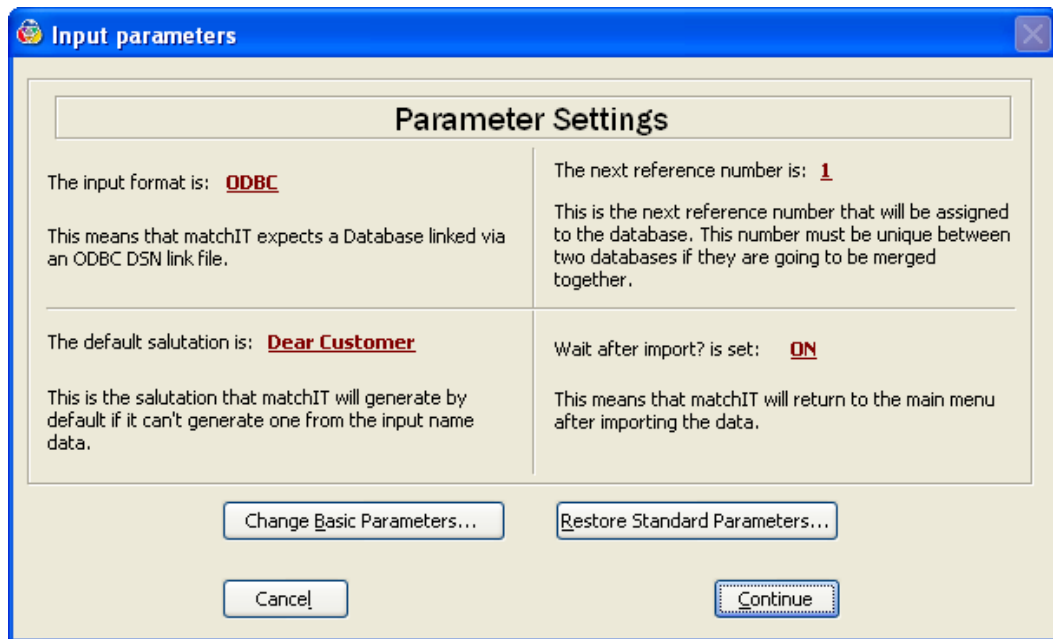
Once you have reviewed and changed the parameters if necessary, select the Continue button to specify the ODBC Connection file that you wish to use. matchIT will then import the data and continue with processing as requested.

Exporting via an ODBC Connection

matchIT can use the same ODBC Connection file to export data back to the data source. To do this, choose an output format of ODBC when specifying the output options. However, take careful note of the created ODB file name and location assigned by the Setup Wizard, so that you can select it for output as you will be prompted for the ODB file. You will then see the following window:
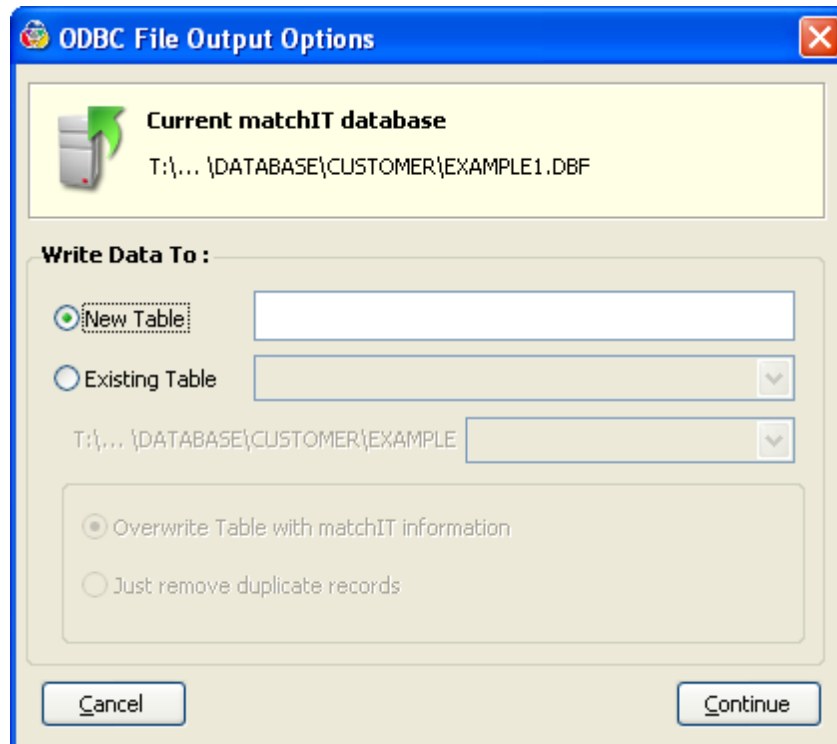
Creating a New Table

To create a new table, choose the "New Table" radio button and enter a table name. The new table will be created in the Data Source specified by the ODBC Connection file.

Overwriting an Existing Table

To overwrite an existing table in the source database, choose the "Existing Table" radio button. Choosing this output option allows you to specify other options which are only available for this type of output. The resultant output window is displayed below.
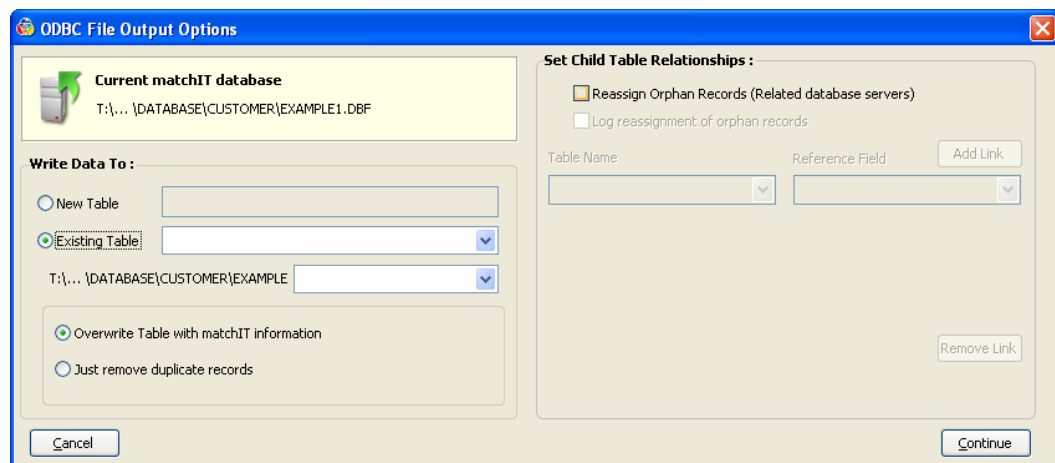
Overwrite Table with matchIT information

This option will overwrite an existing table with information from the matchIT Main File. The type of output will depend on what output option was chosen (unique records, matched records, or flagged records) and the field layout will be determined by the layout option(s) chosen by the user.

Just remove duplicate records

This will remove the duplicate records that you flagged within matchIT from the existing table that you selected.

Both the "Overwrite Table with matchIT information" and the "Just remove duplicate records" options can cause relational data to become corrupt if careful steps are not taken. matchIT has integrated "Child Table Relationships" to help cope with this task.

Reassign Orphaned Records

This functionality is used when working with parent and child tables that are related via a specific reference field. Generally the parent table will be (or should be) a unique list of names (or companies) and addresses. The child table will often be transactional data, linked to the parent table via the reference field. Each parent record will relate to any number of child records in a one-to-many relationship. matchIT is commonly used to dedupe a parent table. This could result in orphaned child records, where parent records are removed and the child records that relate to the removed records are left orphaned (where the references in these child records no-longer relate to references that exist in the parent table).

The "Reassign Orphaned Records" functionality enables users to specify any related child tables during the output step (when writing the deduped parent data back). It will then re-populate the references of the records in the child table(s) with new references. The new references will be those of the records that the original parent records were deleted against.

To use this functionality in matchIT, there must be a parent table and a child table. The child records must contain references linking them to parent records. Parent Data should be imported via an ODBC connection, and upon doing so, it is essential that the reference field linking the parent and child tables together is labeled as 'unique_ref' during the Setup Wizard stage.

The following example illustrates this…

PARENT TABLE

FRED WATERS, 1 High Street, Leatherhead ID110

MR F WATERS,   1 High St, LEATHERHEAD ID103

CHILD TABLE

F WATERS Transaction 31763 ID103 - changed to ID110

F WATERS Transaction 332140 ID103 - changed to ID110

In this example, all F WATERS records in the child table became orphaned, as their parent record had been removed in the parent table. The references for the child records have now been re-

assigned to the new parent record. So 'MR F WATERS' has been removed and all transactions in the child data that related to this record, now relate to 'FRED WATERS' instead.

After importing the parent data and flagging the duplicates, you can write the data back to the source database using the Output To File screen, Output Flagged Records or Output Matched Records. You can select the required fields as normal for the type of output that you are using.

Make sure that the same ODBC connection used to import the data is selected as the output destination, and select ODBC as the output format.

If you select ODBC as the output format, matchIT will display the 'ODBC File Output Options' screen:
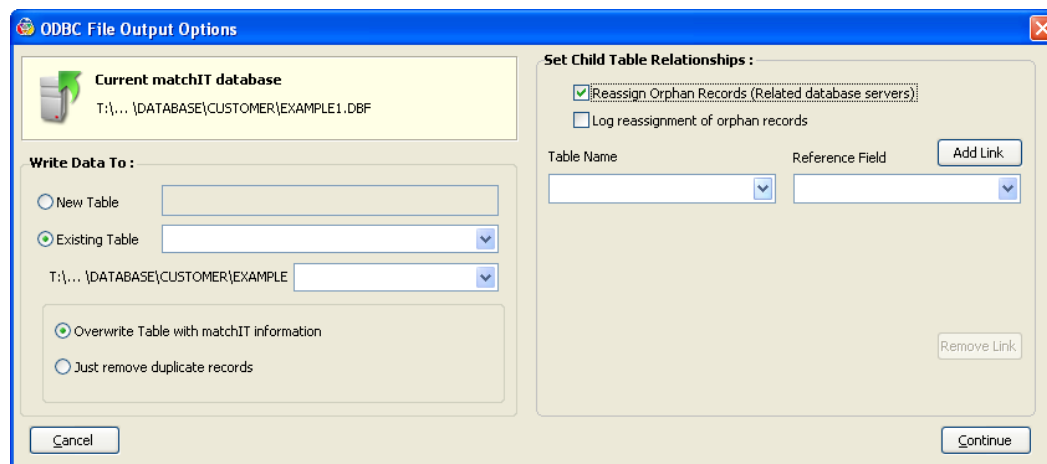
The 'Reassign' functionality is only relevant when overwriting the source table, so from the 'Existing Table' drop-down, select the original parent data table. Next, select 'Reassign Orphan Records' on the right-hand side of the dialog. The child table and the corresponding reference field can then be specified, linking the child records to the parent records. Click on 'Continue' to export the cleaned parent data and re-populate the references in the child table where necessary.

Non Mailable Records

The 'Non Mailable Records' functionality is designed to search for, flag and report specifically on records containing 'mail exclusion' words. The user can add, remove or edit mail exclusion words as required. In addition, mail exclusion words can be grouped together where they represent similar exclusion types. For example, "HMP" and "Prison" can both be added to the same group, as could "Goneaway" or "Moved".

matchIT performs the search for mail exclusions during import, or when re-generating keys on a pre–imported DBF file. All input fields are searched.

When mail exclusion words are found in a file, they are listed and counts of each are displayed in matchIT's standard Data Summary report.

The user can browse the records containing mail exclusions by the type of exclusion, from matchIT's 'Import Details' dialog. So, using the previous example, the user could choose to browse all 'Prison' exclusions, which may display records containing words such as 'HMP', 'Prison', etc.

Adding/Removing Mail Exclusion words

All mail exclusion words must be added to matchIT's mail exclusions table; 'mail_exclusions.dbf'(located in the matchIT's DAT sub-folder). The easiest way to do this is to browse the table via **View>View Table**, from matchIT's main menu. Upon selecting mail exclusions table, it will be displayed on screen and can be edited as required. Of the five fields displayed, each field (except the 'COUNTS' field) must be populated by the user when a new mail exclusion word is added. A description of each field is as follows:

- **EXCLUSION**

  Contains the actual mail exclusion word that matchIT will search for in the input data. All entries here must be upper case. Limited to 10 characters.

- **DESCRIPTIO**

  Contains a description specific to the exclusion word. This will be displayed in the Data Summary report. Limited to 40 characters.

- **XTYPE**

  Contains a code, which will be copied into exclusion records within the working DBF file. Limited to 2 characters. Must be unique to each group/type of mail exclusions (not each individual exclusion word).

- **TYPE_DESC**

  Contains a description of the exclusion group/type. All mail exclusion words in the same group, with the same XTYPE field value should have the same TYPE_DESC. This description will ultimately be displayed in the 'Import Details' dialog, where the user chooses the group/type of mail exclusions they wish to browse. Limited to 20 characters.

- **COUNTS**

  This field is populated by matchIT. It should not be modified by the user.

To add a new field to the mail exclusions table, select **Table>Append New Record**. To remove an exclusion word, simply delete the record from the table by clicking on the small deletion cell to the left of the record.

After updating the mail exclusions table as require, the mail exclusion words must also be added to (or removed from, depending on what you are doing) matchIT's standard Names and Words table via **Job/Setup>Names and Words**. They should be categorized as type 'Exclusion'.

Enabling/Preventing Mail Exclusion in New Data Files

matchIT will only ever scan for mail exclusions on a file, if the file contains a field labeled **MAIL_FLAG**. It is therefore essential to add this field during the Setup Wizard stage, before the import begins. This field is populated with the XTYPE codes from the 'Mail_Exclusions' table, where records contain mail exclusion words. It should therefore be added with a width of at least 20, to cater for records that may contain multiple mail exclusion words.

Where mail exclusion scanning is not required, it is not necessary to add this field.

Reporting of Mail Exclusions

Mail exclusions are displayed on the Data Summary report. The report shows the type of mail exclusion word present in the mail exclusions table, as well as the corresponding exclusion code, count of occurrences in the work file, and percentage of records in the file containing the exclusion word.

If the file that is being reported on is not the last file that had keys generated, matchIT will re-count the mail exclusions. Therefore the report may take a little longer to produce than usual.

If a MAIL_FLAG field is not present in the dbf that is currently open, the Mail Exclusions section will not appear in the report at all.

The Mail Exclusions area of the Data Summary report is displayed below:

Mail Exclusions			
Description	Code	Number of Records Containing Exclusion Word	Percentage of Records Containing Exclusion Word
British Forces Postal Office	BF	2	1.60%
Care Of	CO	1	0.80%
Her Majesty's Prison	PR	3	2.40%
Other Exclusions	OT	1	0.80%

NB: A maximum of 12 mail exclusion words will be displayed above. Please contact helpIT systems if your mail exclusions table exceeds 12 entries.

Browsing Records that contain Mail Exclusions

After importing a file of re-generating keys in a file, any records that are found to contain mail exclusions can be browsed via **Import>Import Results>View Records By Category**.

To browse a specific category of mail exclusion, tick the 'Use Mail Exclusions' tick-box in the bottom right-hand area of the screen. A drop-down menu will now be enabled, which is populated with each different mail exclusion type found in the mail exclusions table, as pictured below:

To browse records containing mail exclusion words falling into any of the categories listed, select the category and choose 'Continue'. You can also modify the output options at the bottom of the main dialog if you wish to send the records to a separate DBF, report on them, or produce a PDF report displaying them.

If a MAIL_FLAG field is not present in the dbf that is currently open, the Mail Exclusions section will not appear in this dialog.

Maintenance

When re-installing matchIT or moving matchIT to a different machine, it is essential that both MAIL_EXCLUSIONS.DBF (located in the matchIT root directory) and the DAT sub-folder are copied to the new installation. This will ensure that any customization carried out will not be lost.

It is also advisable to regularly make backup copies of these files, in case the matchIT installation is accidentally removed or overwritten.

US Address and Zip code Validation

Any version of matchIT can have the US Address and Zip code validation module, **addressIT**, integrated with it, using the addressing engine supplied by Datatech. The standard evaluation configuration does not include **addressIT**, but a user can request an evaluation – we will then send the evaluator a CD containing the United States Postal Service Address File (USPS).
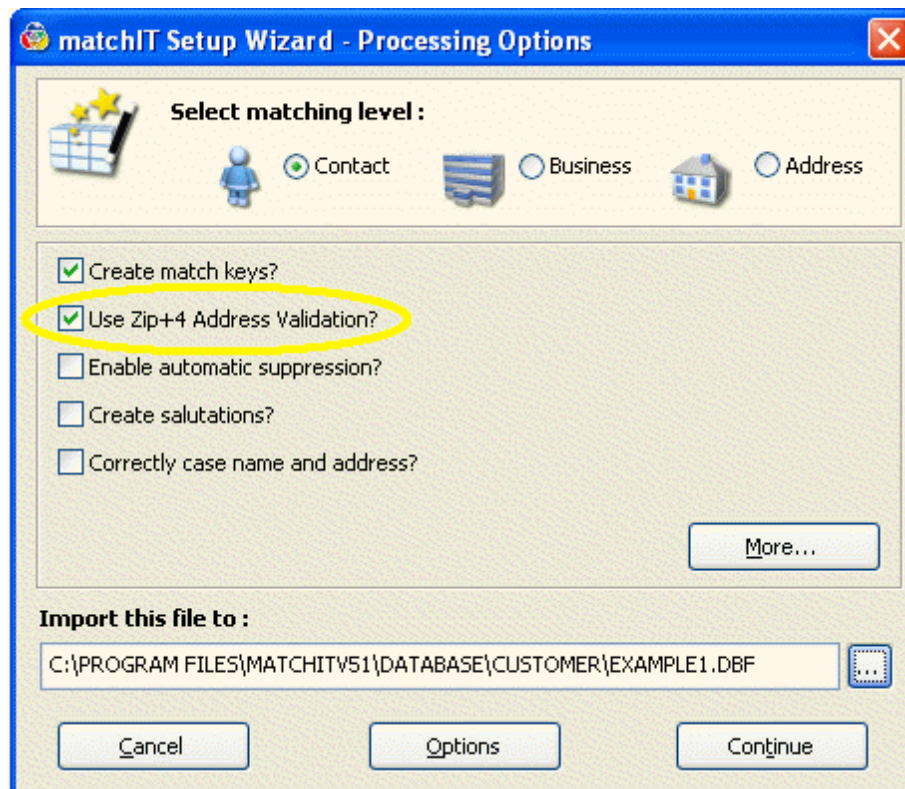
As with all of matchIT's operations, the data must first be imported into a DBF file. The Single or Two File Wizards on the left pane allow the user to do this. The same rules for naming fields, as described in the Getting Started Guide, should be used. Once the data file has been imported via the wizard the user needs to tick the **Use Zip+4 Address Validation** option in the Processing Options window, as shown above. The **Address Enhancement Wizard** will then be displayed after selecting **Continue**.

Address Enhancement Wizard

The **Address Enhancement Wizard** will guide users through the key areas when preparing a file for Zip+4 Enhancement. Each step & the options available will be described in subsequent sections.

Step 1: Configuration & Settings
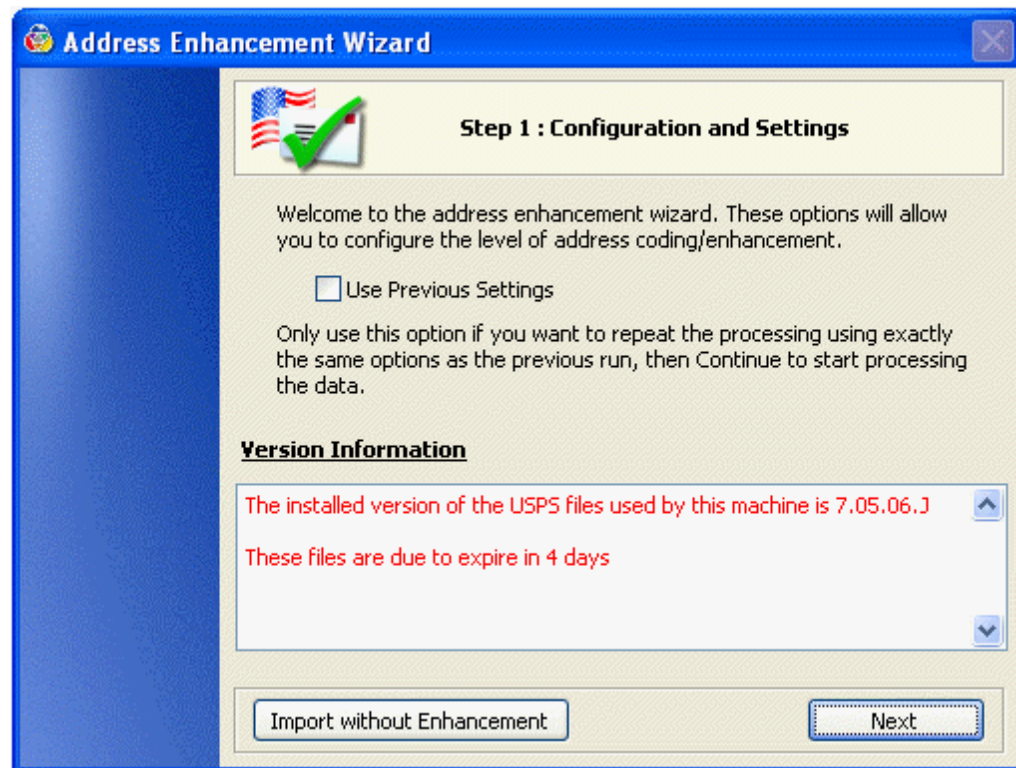
The **Configuration and Settings** window provides the user with several options to determine how they would like to begin enhancing their file:

- **Use Previous Settings** tells the matchIT to use the Zip+4 enhancement options from the previous run.

- **Import without Enhancement** will import/create the matchIT DBF for a source file, without performing any Zip+4 enhancement.

- **Next** will allow the user to go through each section of the Address Enhancement Wizard so that they can specify the options they would like to use on a particular file.

The *Version Information* section displays the expiry date for the **addressIT** data file. This expiry date can also be found in the **Addressing** tab of the matchIT options window. For this run through please select the **Next** button, as it will allow us to walk through each of the **Address Enhancement Wizard** steps.

Step 2: Headline Data Options
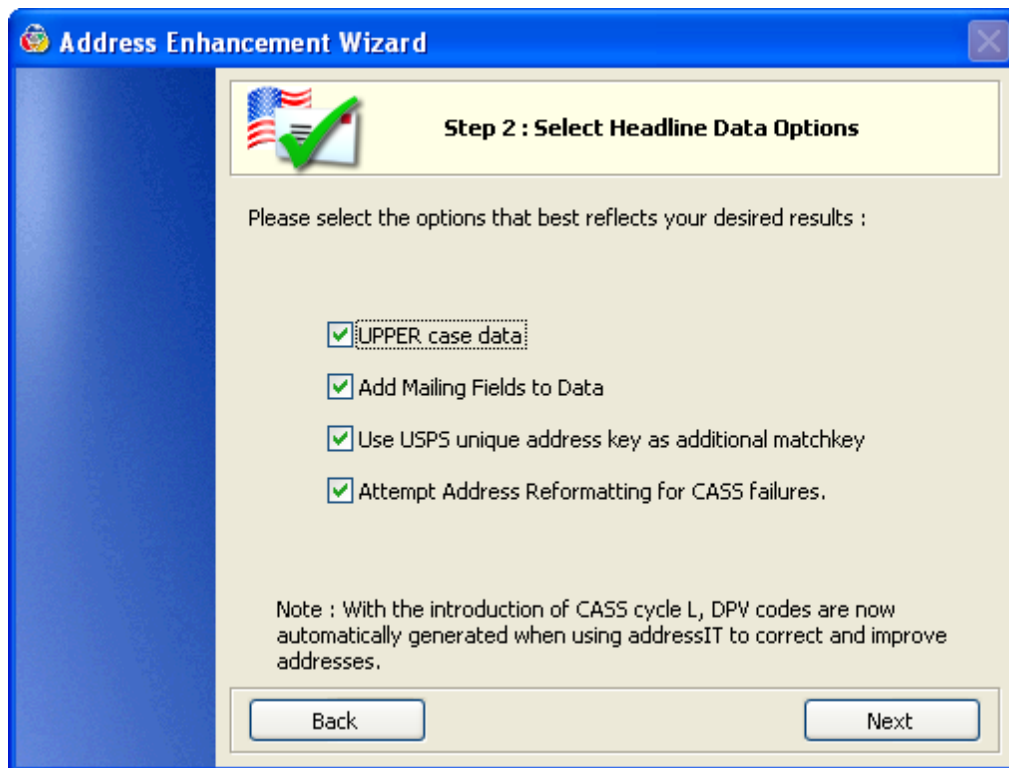
In the **Headline Data Options** window, the user will have the ability to determine which initial options they would like to apply to their data set. The list below will detail each of the options found in step 2 of the **Address Enhancement Wizard**. Some of these options will have a direct effect on subsequent processing, so it is a good idea to read the tooltips provided when hovering the cursor over any of the Headline Data options.

- **UPPER case data** is a specific option that is more likely to be used if a formatted mailing piece is being produced.

- The **Add Mailing Fields to Data** option allows a user to add fields useful in preparing an output suitable for mailing, such as Carrier Route Codes, Line of Travel info or Delivery Point barcodes. If this option is not selected, the **Address Enhancement Wizard** will not display **Step 4: Mailroom**.

- **Use USPS unique address key as additional matchkey** generates a unique key from the returned Zip+4 coding which is useful in enhancing the matching keys and results.

- **Attempt Address Reformatting for CASS failures** will use simple checking to see if the format of the address coding failures can be improved, then will format the address accordingly.

In this walk through, make sure to at least select the **Add Mailing Fields to Data** so that step 4 of the **Address Enhancement Wizard** is not skipped. After selecting the Headline Data options to use, please select the **Next** button to continue to Step 3.

Step 4: Mailroom

The options within the **Step 4: Mailroom** window will only be displayed by the **Address Enhancement Wizard** if the user selected to **Add Mailing Fields to Data** (in *Step 2* of the **Address Enhancement Wizard**).

- To add a Carrier field to the data set, a user can select the **Add Carrier Route Codes**.

- Choosing the **Add Delivery Point Barcodes** option will result in the addition of a field labeled Dpbar.

- Selecting **Add Line of Travel Information** will create a field labeled Lot within the users database.

Once the appropriate options have been selected, the **Finish** button will close the **Address Enhancement Wizard** and allow matchIT to carry on with the importation of the source data.

Address Enhancement Summary
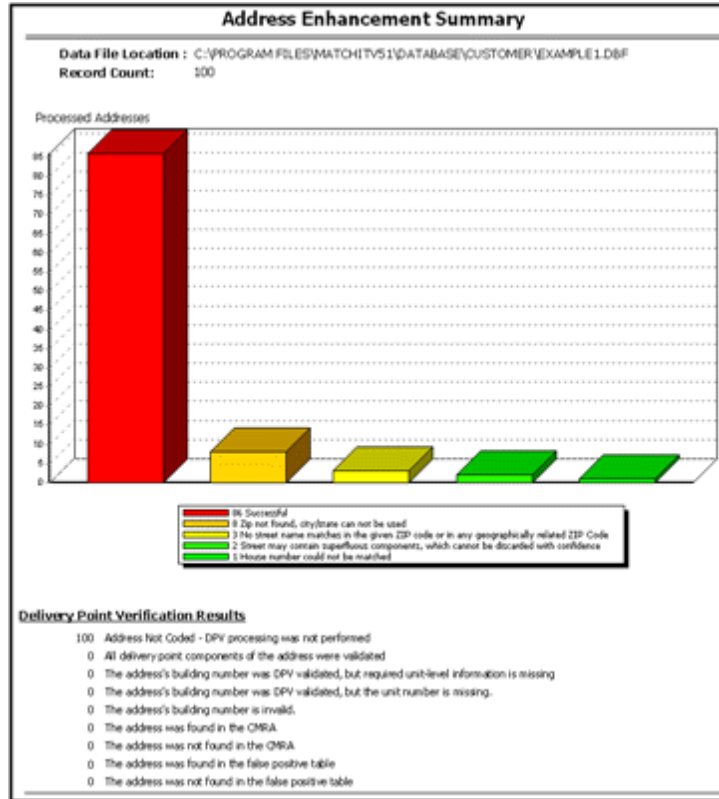
The **Address Enhancement Summary** is a simplified report that provides the user with a quick overview of the address enhancement results, as well as the details for the **Delivery Point Validation** (DPV) option.

Exact figures for the address enhancement graph can be found directly beneath the graph itself. It will display how many records were coded as a success, as well as how many records were coded as failures. Failures will be organized into categories that are based on the reason for the failure.
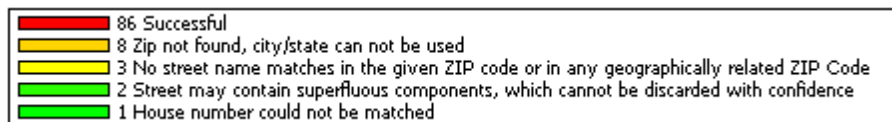
Specific **Delivery Point Validation** (DPV) information can be found in the **Delivery Point Verification Results** section of the report. This section will display the results for the Delivery Point Validation of the addresses found within the users source file.

This reports print toolbar will work in the same way as the other print toolbars. Please use the icon to close the report preview and continue on.

Verifying the addressIT Results
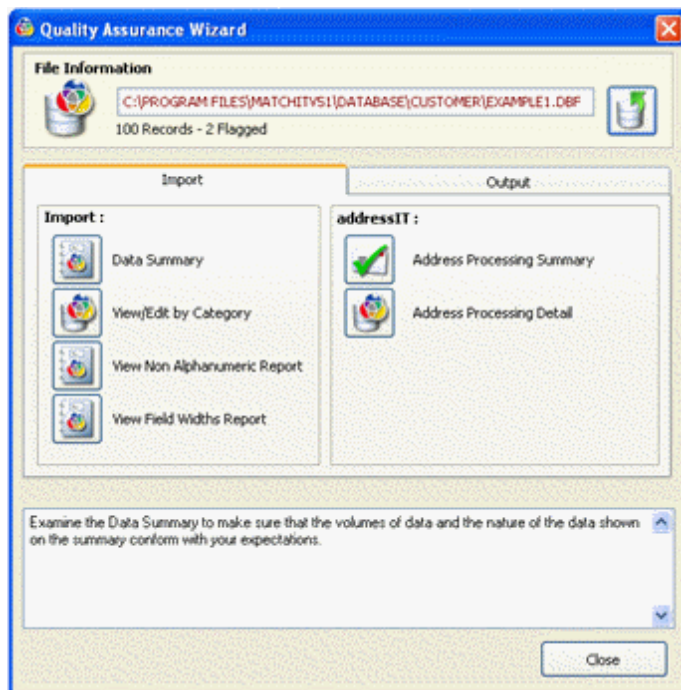
When a file has been passed through **addressIT**, the **addressIT** section of the Quality Assurance Wizard will become active. This section will provide the tools necessary to Q/A check some, or all of **addressIT**'s results.

- The **Address Processing Summary** will allow the user to either preview, print, file or PDF the **Address Enhancement Summary** previously discussed.

- When using the **Address Processing Detail** option to view records, matchIT will allow the user to view either address successes or failures based on criteria selected by the user. (See Address Processing Detail below for more details)

The **Address Processing Detail** option is the most powerful quality assurance tools for **addressIT**. The next section will go over some of the more common **Address Processing Detail** options and functionality.

Address Processing Detail
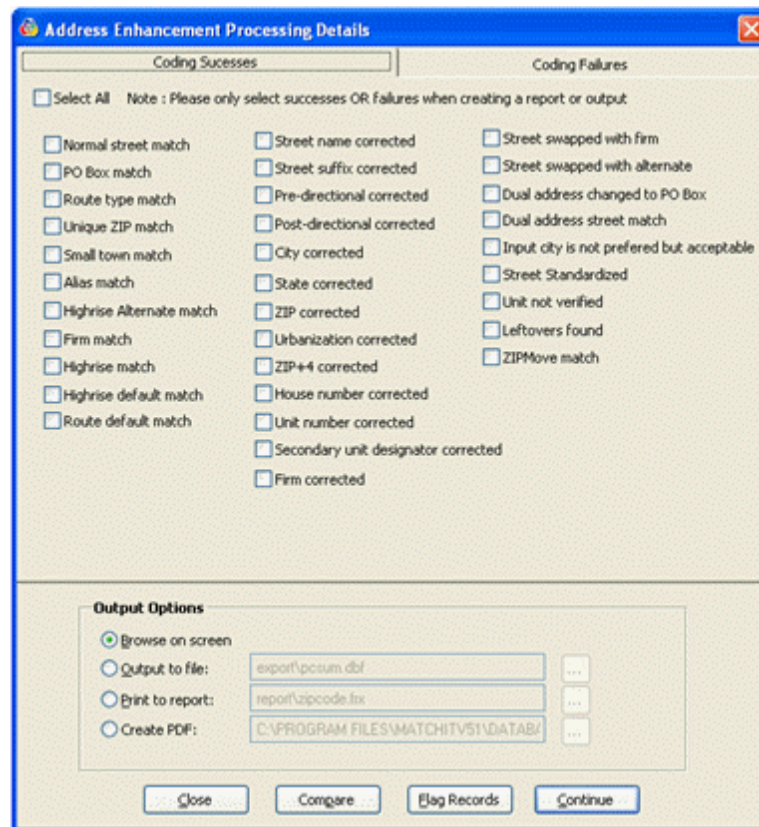
The **Address Processing Detail** option works much the same way as the **View/Edit by Category** option. A user can select one or more criteria for viewing **addressIT** successes or failures. Both the **Coding Successes** and the **Coding Failures** tabs offer an option to **Select All** records. matchIT, however, will not allow the user to Select All for both **Coding Successes** and **Coding Failures**.

In the lower half of the window, the **Output Options** are displayed. These options allow the user to perform several tasks that are based on the coding criteria specified by the user.

- Using the **Browse on screen** option allows the user to compare the input and corrected data in a side by side view.

- Selecting the **Output to file** option will create a separate DBF file containing records specified by the user's selected criteria.

- The **Print to report** feature will send a report to a user specified printer. This report is based on the **addressIT** criteria chosen by the user.

- If a PDF version of the report, mentioned above, is needed – the user can select the **Create PDF** option. They can then specify where they would like the PDF to be created/saved.

The **Continue** button allows the user to carry on with either of the options they have selected, while the **Compare** button allows the user to view data side by side – regardless of the option selected. Choosing **Flag Records** will remove records that correspond with the criteria selected by the user. The following screen shot and descriptions will discuss the **Compare** option provided in the **Address Processing Detail** window.

The window above displays a side by side comparison of the original data and the **addressIT** corrected data. The records displayed will depend on the success or failure criteria specified by the user.

matchIT does allow the user to modify either the original data or the corrected data. Modifying the data, of failed records, can help increase the chances that they will be CASS certified when the records are output in a source file. Once the Close button has been selected, the changes will be saved and the user will be returned to the main **Address Processing Detail** window. Returning to the main **Address Processing Detail** window will allow the user to carry on with any other processing of CASS successes or failures. Selecting the Close button will return the user to the **Quality Assurance Wizard**.

The next section will discuss the use of **addressIT** on an output file. From the **Quality Assurance Wizard** window, please select the **Output** tab. From the **Output** tab, choose the **Output to File** button. The **Produce Output** window should now be displayed.

Utilizing addressIT for an Output File
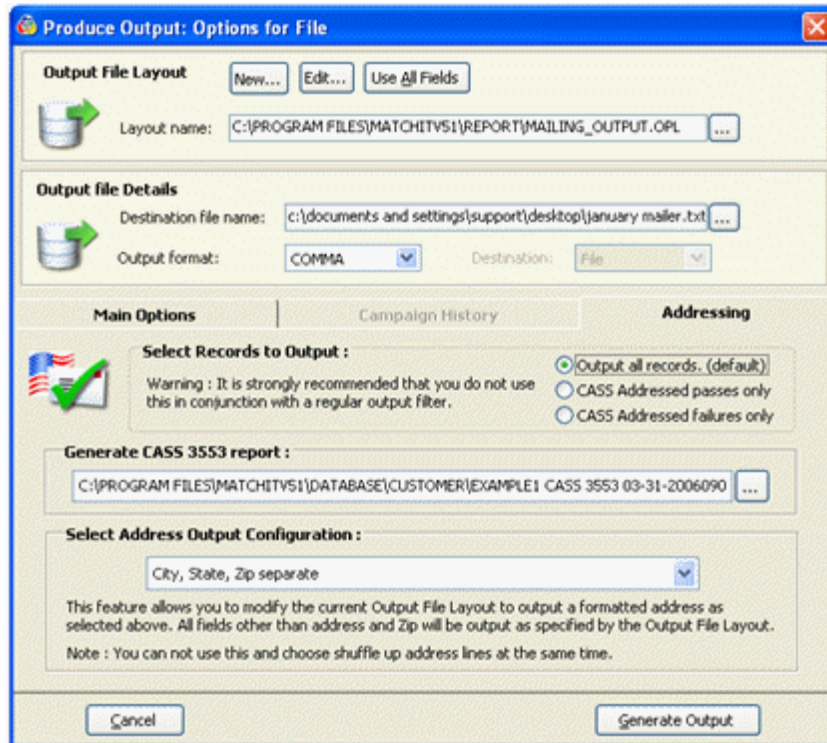
The **Addressing** tab of the **Produce Output** window will become active when the user tries to output a file that has been processed by **addressIT**. These new **Addressing** options will be available in addition to the **Main Options** for outputting a file. The inclusion of these addressing options will allow the user to output subsets of their file depending on the CASS status of a record, while providing the ability to format the address lines and generate a 3553 CASS report for the resultant file. For more details on these options, please see the details below.

- The **Select Records to Output** section allows a user to determine if they would like to output all records, CASS passes, or CASS failures. This option will place a filter on the source file, so as to only output the records required by the user.

- To **Generate CASS 3553** report the user only needs to specify the path of where they would like the report to be saved. Once the output file has been generated, the CASS report will be run on that file. The user can then find a TXT and PDF version of the report saved to the location they specified. If **addressIT** was being used for a mailing piece, the 3553 CASS report will need to be supplied to the United States Postal Service (USPS).

- To format the address lines of an output file, the user can select the drop down option located in the **Select Address Output Configuration** section. The drop down contains several options for the most common address formats.

  - City, State, Zip separate

  - City & State combined, Zip separate

  - City separate, State & Zip combined

  - City, State, Zip combined

- Do not modify the Output File Layout

Selecting any of the options in the **Addressing** tab, can and will modify the source file layout. If the user would like to maintain a layout that has been specified, they will need to set the **Select Records to Output** option to **Output all records** and the **Select Address Output Configuration** option to **Do not modify the Output File Layout**.

suppressIT

Using suppressIT with One Mailing File

Importing Records

The easiest and safest way to use the **suppressIT** module is to use the Setup Wizard to set up any new job, as with the rest of the matchIT suite. For instructions on how to use the Setup Wizard, please refer to the Getting Started Guide.

At the end of the Setup Wizard, you are presented with tick boxes to Create Match Keys, Create Salutations, Correctly Case Data, Use Zip+4 Address Validation, or Enable Automatic Suppression.
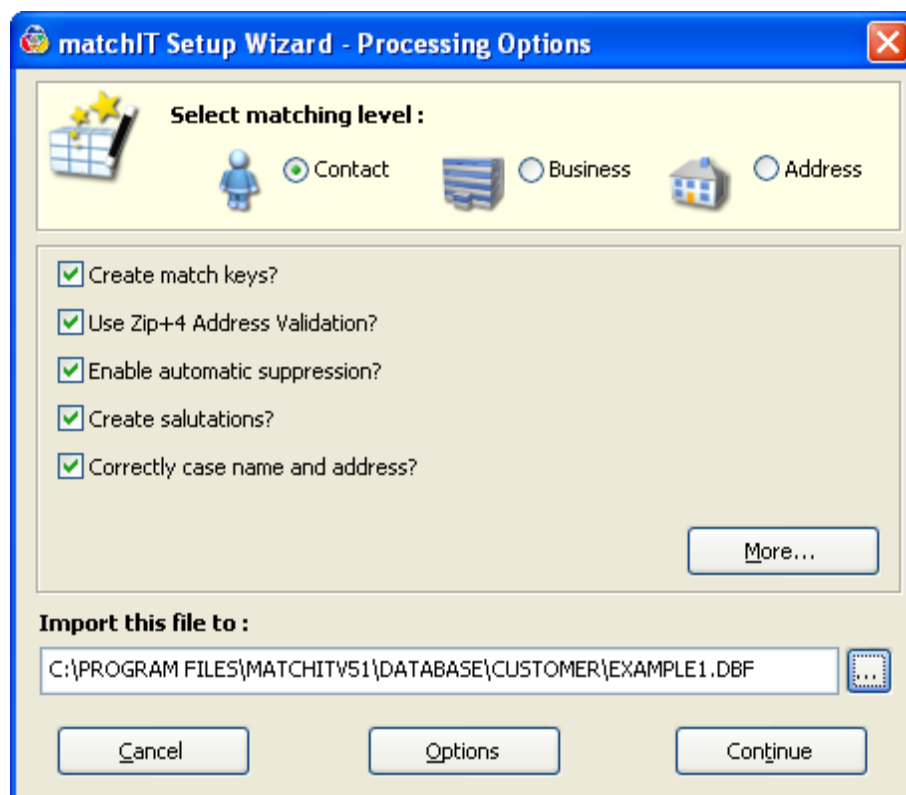
Setting up Automatic Suppressions

Clicking the Options button will allow you to set the Suppression options (Suppression tab). From here you can select the suppression files you wish to use. To set the files, you need to tick the relevant box.

Alternatively, if you choose Find Suppressions from the Suppressions menu you will be presented with a similar screen. This should only be done after all internal duplicates have been removed from the file.
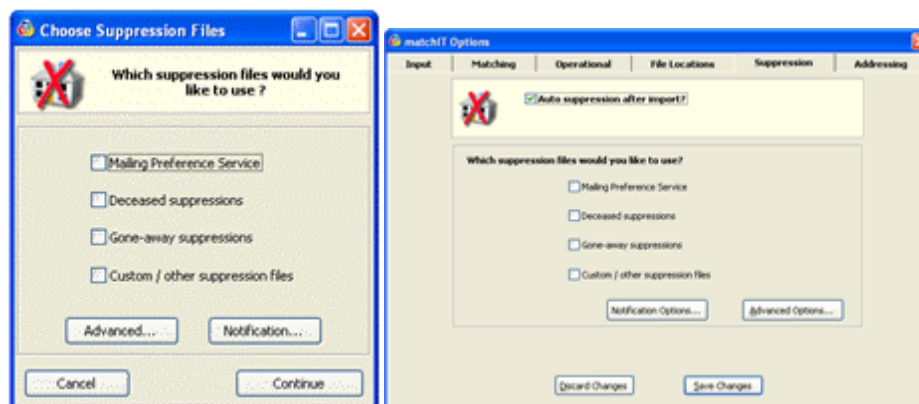
You can set the matching options and file locations for each file, by clicking the **Advanced** button.

As explained above, for initial installation, you need to check the Advanced Options, to ensure the correct files and matching options are being used. You only need to do this on initial implementation, unless you need to add more suppression files at a later date. For directions on setting up the Suppression files, see

The Matching/Suppression Process

When you have set all the **suppressIT** options, you are now ready to import the data. You will see the only option available is **Continue** or **Find Matches**, depending upon if you selected the **More…** option. This will create the matchIT database for your source file, add matchIT's fields to this database, copy the records into this database, and then find the internal duplicates (within the source file). When it has found the internal matches, it will automatically find the suppression matches – dependent on what suppression files you have specified.

As stated above, if you check the **Enable auto-suppression** box & then **More…**, you will see the only option available will be **Find Matches**. If you don't want to import the file yet, or want to **Import Only** (e.g. if you don't want to use the default rules for finding matches, or you want to use a different level of matching for the suppression files), you should uncheck the **Enable auto-suppression** box to enable the other options as usual. You can then run Import and Find Matches from the menus or when prompted, and run the Suppression processing from the Suppression menu, or when prompted after Find Matches.

Suppression Results

When the Find Suppressions step has finished, matchIT displays the **suppressIT** Results. From this screen you can run the Suppression Summary Report and the Suppression Matching Report. You can also Remove Suppressed Records.
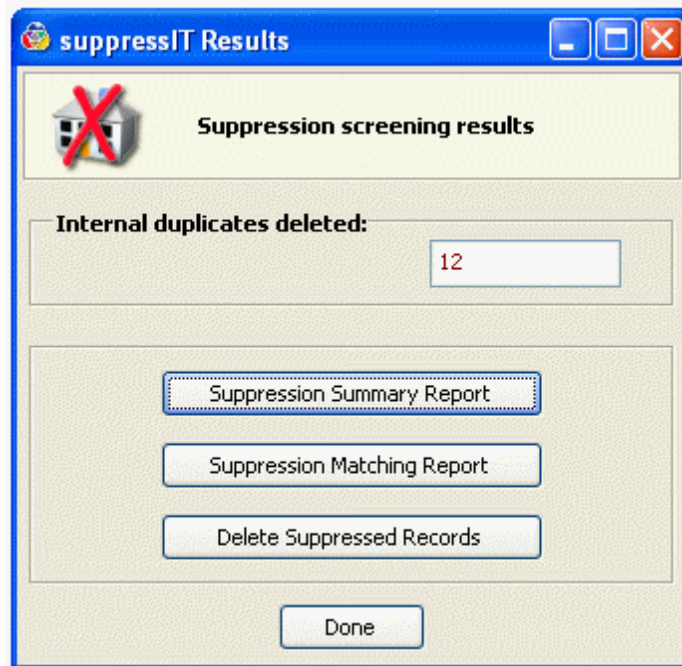
Suppression Summary Report

This report can either be run from the **suppressIT** Results screen or from the Suppression menu. This report summarizes the suppressions found in three ways.

- The Number of Potential Suppressions is the number of records on the Client Database that will be suppressed. Please note that this number will not necessarily tally with the totals in the other sections. One Client record could match with more than one suppression record, but would still only be counted as one potential suppression.

- The Suppressions Found by File section lists the total number of suppressions found in each suppression file. These figures represent the total number of Client records that had matches in each suppression file. One Client record with more than one match in a particular suppression file would be counted once for that suppression file.

- The Suppressions Found by Score Range section lists the total number of matches found in each Matching Score Range. These figures represent all of the matches found. So, if one Client record had two matches in one suppression file and one match in another suppression file, it would be counted three times in these figures. This section is useful for understanding the results you are getting at different Matching Score levels.

Suppression Matching Report

This report can either be run from the **suppressIT** Results screen or by choosing View Suppressions from the Suppression menu. If run from the **suppressIT** Results screen, the matches

will be grouped in Sets and sorted by Score. If run from the menu, the View Suppressions Options screen will be displayed.
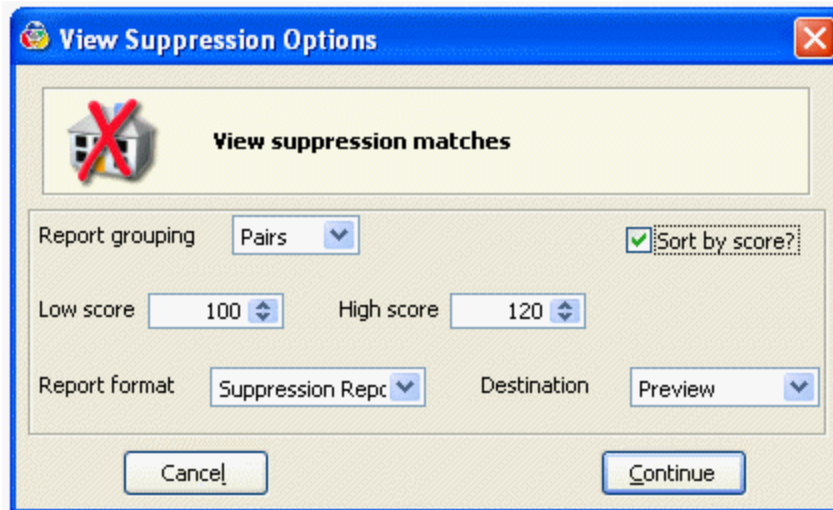
Report grouping

Matching records can be grouped in Pairs of one client record and one suppression or Sets where all suppression records matching the same client record are displayed together.

Sort by score

If this box is ticked the report will be sorted by matching score, otherwise it will be sorted by name.

Low score and High score

The range of matching scores for which the results are to be displayed. Only matches within this range will appear on the report.

Report format

The report format will default to Suppression Report. Alternatively, you can choose your own report layout from here.

Destination

Preview for a print preview, Printer to print the report or File Preview to produce output to a file.

Continue

Click **Continue** to confirm your choices and create the report.

Delete Suppressed Records

This can either be run from the **suppressIT** Results screen or by choosing Remove Suppressions from the Suppression menu.

Minimum Score

Records with a matching score equal to or greater than this value will be flagged for exclusion from the output file.

Continue (accept cost)

This will remove all records meeting the Minimum Score above.

Warning: You will be charged as soon as this button has been pressed. This charge cannot be undone.

If you have the option Notify Suppression Hits By Email turned on, they are sent via an Internet connection to your supplier. If you don't have an internet connection available on the PC where you are running **suppressIT**, you will be asked to email a log file to your supplier.

Cancel

Cancel the deletion process. You will not be charged.

You can incrementally remove score ranges if you wish (starting at a high score and getting lower- obviously, the lower the score, the more records will be removed). You will only be charged for the new records removed each time. *Remove Suppressions* is incremental – you will not be charged twice for the suppression, unless you "undelete" the records. (**We strongly recommend that you do NOT undelete records**).

Also, we suggest that you flag your internal matches first – this way, some of the potentially suppressed records will be removed before you come to suppress them. Since they are already removed, you won't be charged for records that have been removed due to internal deduplication.

If you do not want to remove all the categories of suppressions that you have found, then you can switch categories off by using the Suppression Options screen (Choose Suppression Files) as illustrated earlier, before you remove the suppressions.

Once you have removed your suppressed records, you can continue using matchIT in the usual manner, for outputting records, etc.

Change of Address Suppressions

As well as supplying a list of names and addresses of those who have moved to a different location, the Equifax Goneaway file can also be used to help locate a large number of people at their new address.

After suppressing a file against the Equifax Goneaway file, select Output Change of Address File from the Suppression menu. This will create a new file listing all of the suppression hits that have new address associated with them.
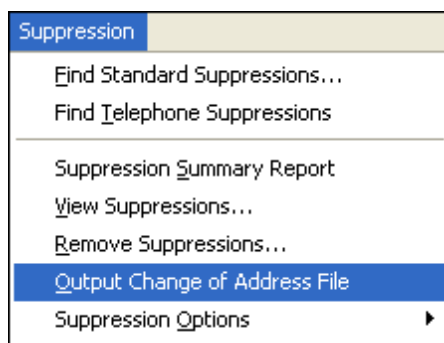
The list created by this function will contain the Unique_Ref values from the records in your file, as well as the Equifax Change of Address reference from the Goneaway records that each of your records were suppressed against. This list can then be sent to Equifax, who will supply back the new addresses associated with your suppressed records.

Please note, not all records on the Goneaway file have a new address associated with them, which is why it is very unlikely that the number of suppression hits will match the number of records in the Change of Address file.

Telephone/Fax Suppression

As well as providing a way of suppressing your file against name and address suppression files, **suppressIT** also provides a facility for suppressing against lists of telephone or fax numbers.

In order to use this functionality, you must first import your data using the Setup Wizard, ensuring that your telephone or fax fields are labeled 'telephone' or 'fax'. Once you have finished importing your data, you can begin the telephone or fax suppression by selecting **Find Telephone Suppressions** from the main **Suppression** menu. This will produce the Choose Telephone Suppression Files dialog, as pictured below.
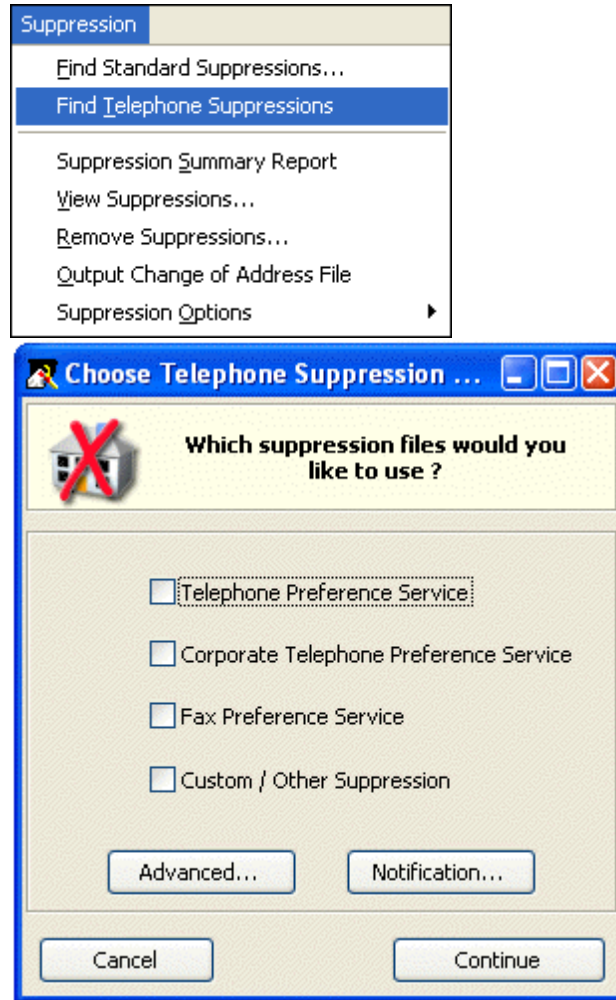
From within the Choose Telephone Suppression Files dialog, you can tick the suppression file(s) that you would like to suppress your data against. After making your selection, you can click on **Continue** to begin the suppression step. Before clicking on **Continue**, you should make sure that you have mapped **suppressIT** to the suppression files that you have chosen to use. To do this, click on **Advanced** to display the Advanced Telephone Suppression Options screen.

Advanced Telephone Suppression Options

From within the Advanced Telephone Suppression Options screen, you are given the facilities to point to 'Add' suppression files available to your PC, no-longer point to 'Remove' suppression files that you no-longer wish to use, or 'Create/Update' suppression files. There are four different areas within this screen. Each area represents a different type of telephone suppression file. These four areas correspond to the four tick-boxes in the Choose Telephone Suppression Files dialog.

Create/Update Suppression Files

From within the Create/Update Suppression Files dialog you can choose to either import suppression data into an existing DBF file, or create a new suppression DBF from scratch. Before making this selection, you must specify the format of your input suppression data file by using the **Suppression File Format** drop-down menu at the top of the dialog.

If you already have an existing suppression DBF file, either select **'Append new suppression records to an existing suppression file'**, to simply add to the data you already have, or choose **'Replace all records in an existing suppression file'** to overwrite the data you already have, with data from a new file. After checking the appropriate radio button, click on **Import Records** to begin importing the new data.

If you do not have a suppression file already set up, you can select **Single field suppression file'** to create and populate a suppression file from a single list of telephone or fax numbers. Alternatively, if your input suppression data file contains multiple fields, select either **'Custom field layout'** to manually name each of the fields in your file, or select **'Pre-defined industry standard field layout'** and select one of a number of industry standard suppression file names, of which **suppressIT** is already familiar with the layout. After making your selection, click on **Create File** to begin the process.

Using suppressIT with the Multiple File Wizard

Setting up the Multiple File Job

Before you enter the Multiple File Wizard, please make sure that you have set up the locations of the suppression files that you want to use. Then set the files up for the Multiple File Wizard as usual.
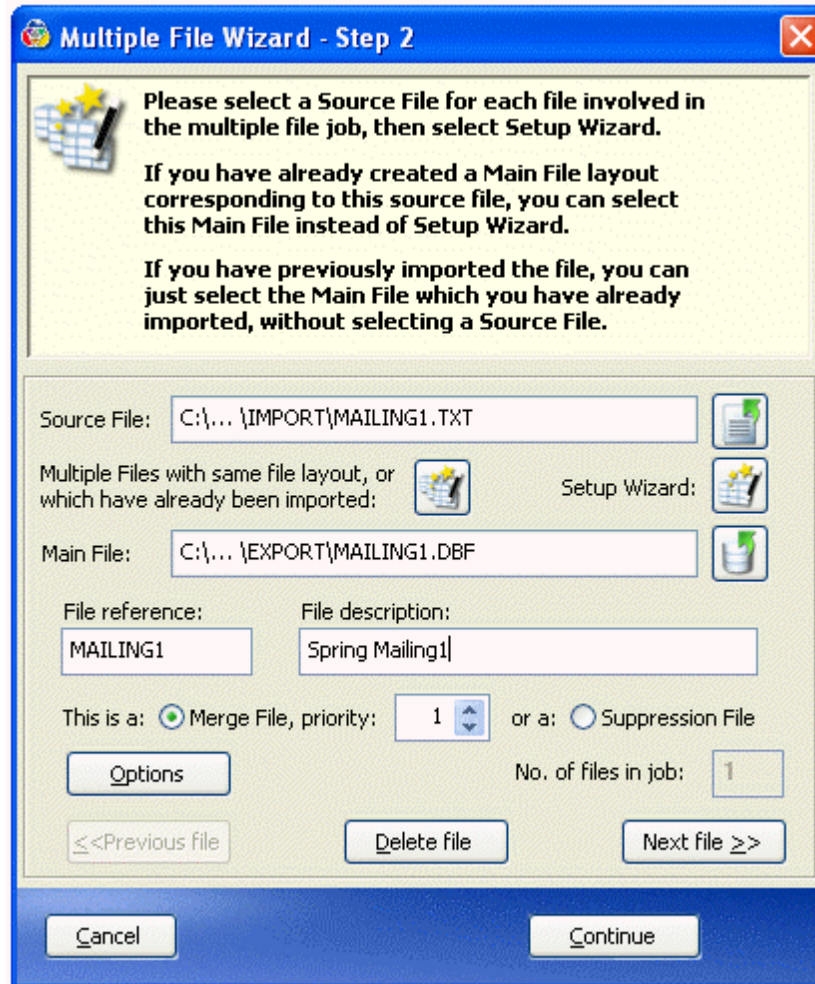
Do not select any of the standard external suppression files within this step – if you have any suppression files which are smaller than or comparable to the total mailing volume, then select them as suppression files here – otherwise, set them up as external suppression files, as described in the next section.

When you specify details of the Master File, select *Options*:

Then select the **Suppression Tab**, tick the box labeled **Automatic Suppression After Import**.
Here you can also change the Suppression Options as usual.
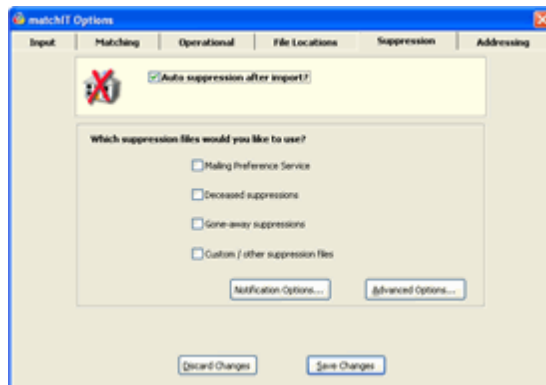
As the final stage in setting up the Multiple File job, **do NOT select to produce the matrix and summary reports**, as you must review and accept the suppression charges (and remove the suppressions) before the matrix report can be produced.
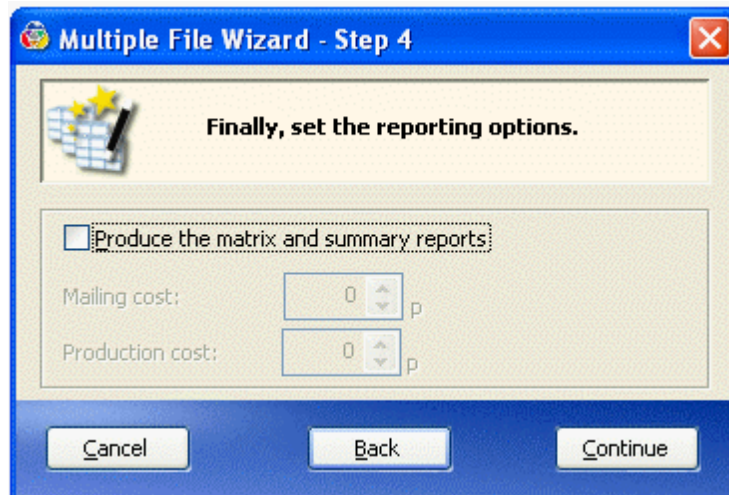
You can then run the job as usual.

Producing Multiple File Reports

When the job has finished, select the **Suppression Summary Report** and then to **Remove Suppressions** from the **Suppression** menu.
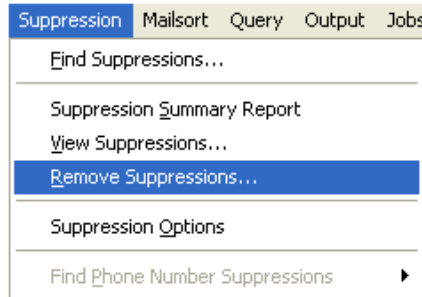
Finally, select to Produce Matrix Reports from the Jobs/Setup menu and select the Job reference that you have just run.
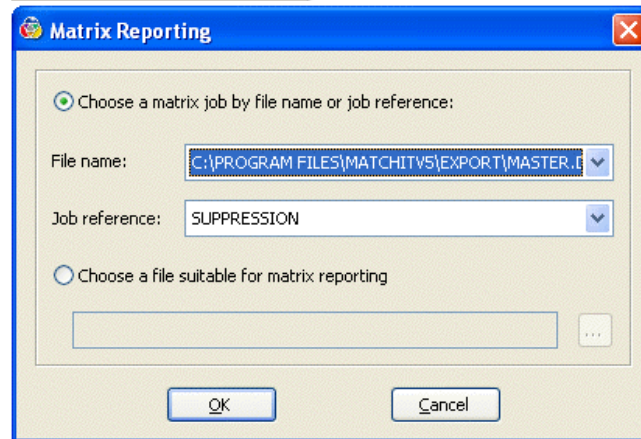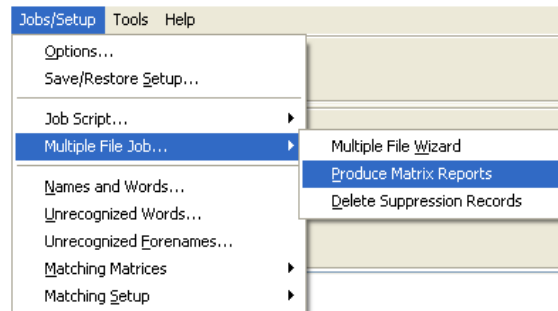
Select to Regenerate the report, enter the costs and the report will then be produced.
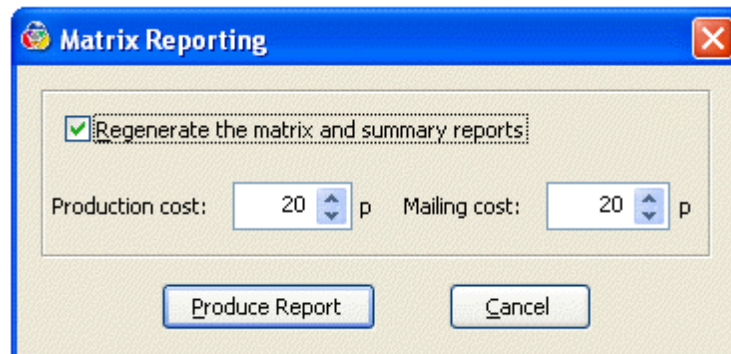
suppressIT Implementation

Suppression File Locations

Before you use **suppressIT** in matchIT, you must set matchIT up to point to the suppression files that you have imported or loaded from the supplied suppression CD. Unless your network is geared up for fast indexed I/O on very large files, we recommend that you place these into a folder on your local machine.
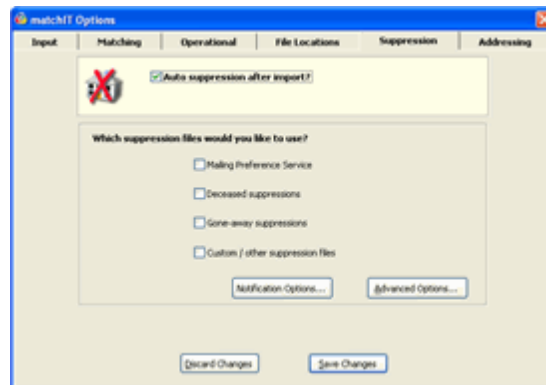
Now open matchIT, and choose **Suppression Options** from the Suppression menu and click on the Advanced button. This will take you into the **Advanced suppressIT Options** screen, which will allow you to set the locations of the suppression files. You can select a suppression file by clicking the **ADD** button.
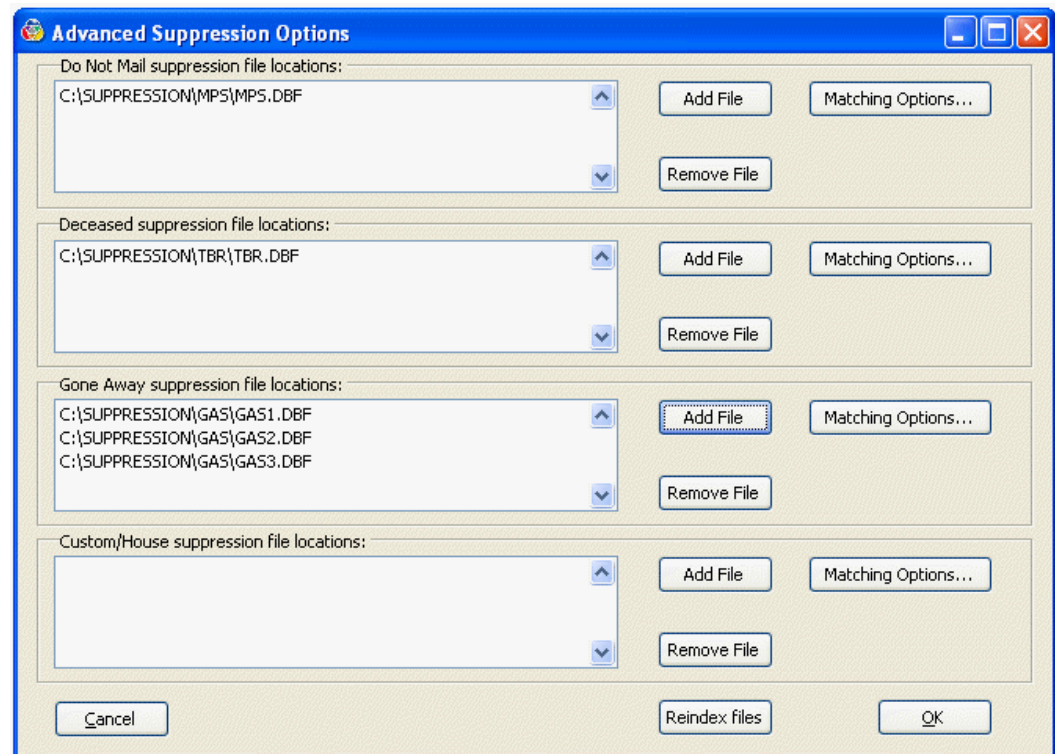
You will need to **Add** each of your suppression files to this screen. Generic terms for the categories of suppression files are being used. These are provided for your convenience. If you

put a file into the wrong category, matchIT will still process the file correctly, but your summary report will not reflect the right categorization.

The **Remove** button allows you to remove it from the list, respectively.

When you have added all your files, you must click on the Matching Options button for each file set, to set up the matching options that will be used when matchIT runs the suppressions for you.

Selecting Matching Criteria

When you have set the files you wish to use, you need to set the Matching Options/Keys. Clicking the Matching Options on each section displays the Suppression Matching Options dialog box.
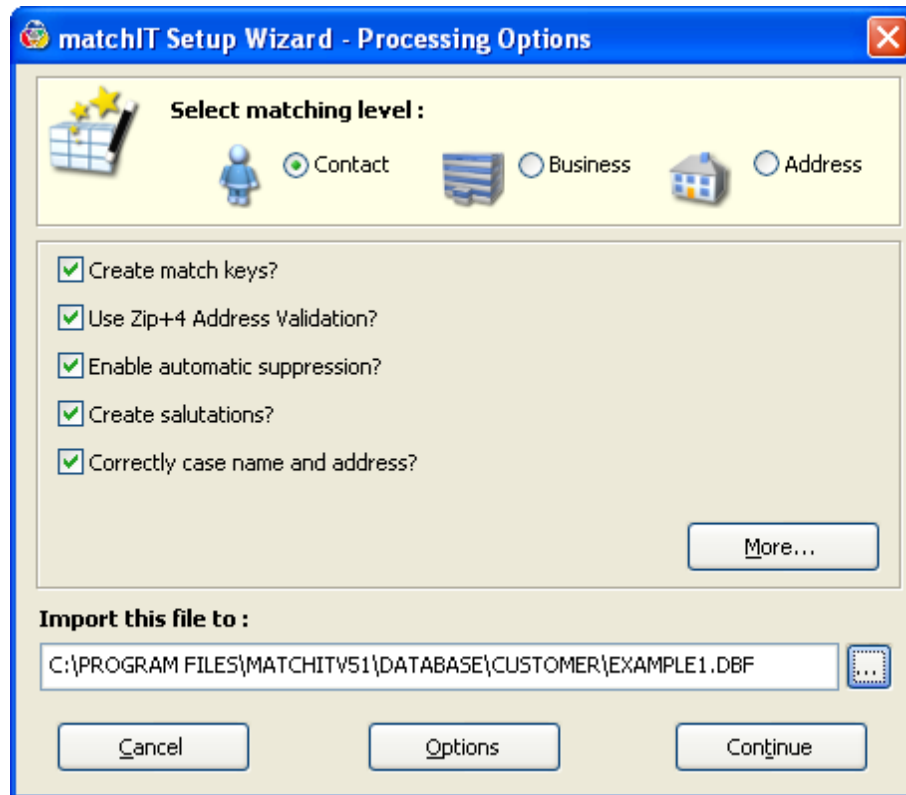
For the standard MPS, Deceased and Gone Away suppression file locations, use the default keys as supplied unless you have a good understanding of the implications of different keys for run times and hit rates (see the Help or User Manual topic "Match Key – Definition". If you're not sure what match keys to use, please call us. The default keys that we supply with **suppressIT** are as follows:

- MPS (Mailing Preference Service) and DDNC (Deceased Do Not Contact)

  ADD_KEY+NAME1

  LEFT(ZIP,5)+SUBSTR(ADD_KEY,5,8)+LEFT(PREMISE,5)+NAME1

You can also specify matching options here: the minimum score to report for each file, the matching level and options for what to do when premise numbers do not match. By default, **suppressIT** is set to match deceased files at individual level and MPS and gone away files at family level, with options set to insist that premise numbers match and do not match if one is empty.

If you have updated the match keys for any of the files, please say Yes when matchIT prompts you to reindex the files, or choose Reindex Files from the **Advanced suppressIT Options** screen, otherwise every time you use **suppressIT** you will have a significant delay while it reindexes the suppression files that you are using.

Company Contact Details

The final part of **suppressIT** you will need to set up before you use it is the Company Contact Details. To access these, choose the **suppressIT** Options from the Suppression menu. You will see a screen like this:
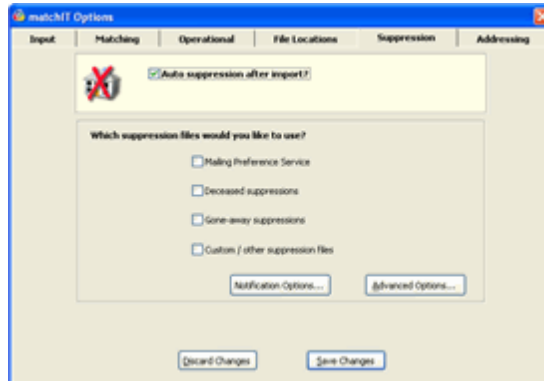
Now, click on the **Notification Options** button, and you will see the Company Contact details screen below:
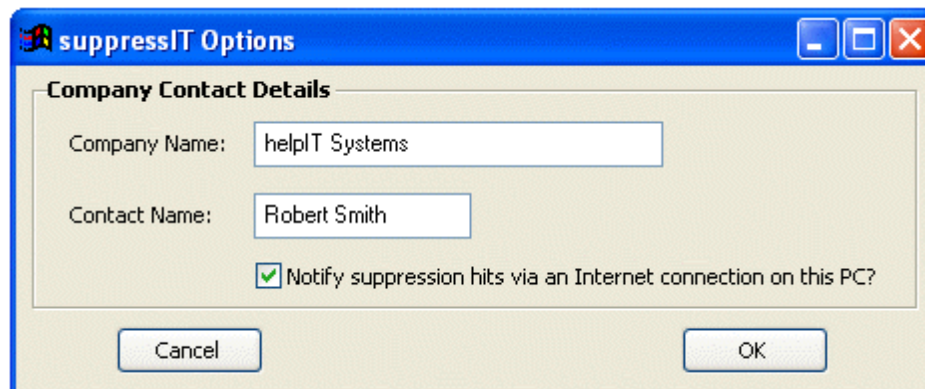
Make sure that the **Notify suppression hits by email** box is ticked (if it isn't, you won't be able to enter the contact information) and then fill in your company name and a contact name. Usually the contact name will be the person who most uses **suppressIT**.

Now click on OK twice, to return to the main menu.

Loading Custom Suppression Files

The first thing you will have to do is to get the suppression files that you want to use with *suppressIT*. Usually, these files will be provided in matchIT-compatible DBF files from helpIT systems. If they are, please skip to the next section, Suppression File Locations. If you are using suppression files not supplied by helpIT systems, please follow these instructions:

- If you received your data in a text file, this means that you will have to import it into matchIT as for any other file that you use with matchIT. If you use the Setup Wizard to specify the layout of the file, select to *Add* a field at the end of the layout and label it *S_ID* with a field width of 2 character.

- If you have already imported the file into matchIT, you will need to modify the structure of your suppression file (from the Jobs/Setup ... Main File Layout menu option), to add the *S_ID* field.

- You must then populate this field with a file identifier. This identifier must be the same for all records in your file, but does not have to be unique across other files.

To do this, you can go into matchIT's system commands (from the Tools menu), and type:

> USE ? (followed by clicking on the Process Command button)

Now choose your suppression file.

Now type:
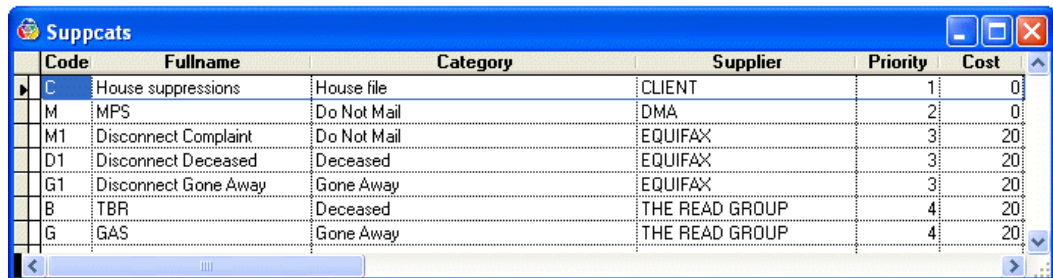
> REPLACE ALL S_ID with "C"

or C1, C2 etc. if you have more than one custom file. The "C" must be upper-cased.

If an imported file will exceed 2Gb after it has been imported into matchIT, you must split it into two or more smaller pieces, each imported into its own matchIT table. If you have multiple matchIT tables for the suppression file, each table that you split the file into should have the same S_ID value.

In terms of performance, it does not matter if you split a file into pieces or not, as the processing time will always be in proportion to the total number of records you are suppressing against.

Next, you must browse the SUPPCATS table in the matchIT directory (which you can do using the View Table option in matchIT's View menu) to add an entry for the code letter that you have used. The supplied table looks like this:

To Add a record, use the Table menu, Append New Record.

The Fullname field is the name that will appear on the reports.

The CODE field corresponds to entries in the S_ID field of your suppression tables. **NOTE: If you have S_ID values that are not present in this table, they will NOT appear in your reports!**

The PRIORITY field indicates the priority of the file. Lower values = higher priorities. Therefore, suppressions will be logged against priority one files in preference to priority two – and so on.

The CATEGORY field just gives some information about the content of the file for the user.

The COST field indicates the price of a suppression (in pence). *suppressIT* will always take the cheaper suppression record in preference to the more expensive one for equal priorities.

Unless you have specific requirements that are different, use the entry for the Mailing Preference Service as a model for your new entry.

Implementation Checklist

Finally, use the following checklist to ensure that you have done everything necessary:

- Imported (or copied) the suppression DBF tables.

- If they are Custom Suppression Files, they all have a fully populated one-character S_ID field and set up entries in the SUPPCATS pricing table for *suppressIT*

- Pointed *suppressIT* to each file location, via the *Advanced suppressIT Options* screen.

- Set up matching keys & options for each file classification on the screen.

- Provided company identification details.

If this has all been done and checked, then you are ready to use *suppressIT*.

Support for Additional Suppression Files

Suppression files, which typically contain several million records, should, be loaded using automated job scripts. This provides the ability for files to be spilt where required, and prevents users from having to remember specific vital steps that are required when setting up a suppression file.

The 'support for additional suppression files' functionality is comprised of one master template job script (including the program set), and five copies of this script, which have been adjusted to cater for the following suppression files: Purity, ESF, Mortascreen, Baby MPS and NCOA. The different copies of the master job script vary based on the structure of the suppression files they were created for, as well as the number of records contained in each suppression file.

When loading a suppression file using it's dedicated script, the user will simply be required to edit the start of the first program script by adding the path, name of the source data file, and the format. The user can then start the script and leave it to run. The length of time it takes to complete is dependent on the size of the file and the performance specification of the PC that matchIT is running on.

When the suppression file load process finishes, the suppression files will be located in the same folder as the source data folder, and will be ready to use in matchIT.

Breakdown of Steps Taken by Job Scripts

The Suppression Load process is made up of three main stages. Each stage, and the corresponding row(s) displayed in the Job Editor screen are explained below.

For further in-depth details relating to any of the steps in the load process, please contact helpIT systems (see page 173).

- **Stage 1 – Split and Append Source Data (first row in the job script)**

Stage 1 is represented by the very first row in the job script. In this row, matchIT uses the SPLIT_SUPPRESSION program to identify the suppression source file (specified by the user), and then split the file into portions. The number of portions depends on the width of the file and number of records. The source data portions are then appended into 'load' tables.

- **Stage 2 – Generate Keys in the 'Load' tables (middle rows in the job script)**

Stage 2 is represented by the second row in the job editor, and further rows depending on how many files the suppression file has been split into. Each row generates keys in and address processes a suppression file portion contained within an intermediate suppression load file.

- **Stage 3 – Create Final Suppression Files (last row in the job script)**

Stage 3 is represented by the last row in the job script. In this row, matchIT creates the final suppression files in the same directory as the source data file. It creates the appropriate indexes for the file(s) and names them appropriately (suffixed with a number if there is more than one suppression table).

Before running a Suppression Load Job Script

1: Create a suppression load directory

The first thing to consider before using a suppression load script is the location of the suppression source data file. This is important, as all processing is carried out in this directory, so it must be a directory on a local drive with plenty of free disk space.

The name of the directory containing the suppression source data file is used by the programs in the script, to determine which suppression file is being loaded. This will ultimately determine which pre-defined 'master structure' file is used to load the data into. It is therefore essential that the directory be named correctly.

All pre-defined job script templates can be found in the "Suppression_Load" directory, which is in matchIT's "Progs" sub-folder. The sub-folders within "Progs\Suppression_Load\" reflect the suppression load directory names that the script would recognize (e.g. the suppression source data could be stored in "D:\Data\Suppressions\Purity" but not "D:\Data\Suppressions\Puritydata").

2: Specify the suppression source file details from within the job script

After creating the suppression load directory and copying the suppression source data file into it, the details of this file must be specified in the job script. To do this, select Automation from the main matchIT screen. From within the Job Maintenance dialog, select the corresponding job script using the 'Job Reference' drop-down and then click on 'Modify Job'. The job script will now be displayed in the Job Editor window.

The suppression source file details will now need to be specified in both programs featured in the script: SPLIT_SUPPRESSION.PRG and CREATE_FINAL_SUPPRESSION_FILES.PRG. These can be edited in turn by double clicking on the program and choosing 'Edit Source File'. When each program is displayed on screen, modify the appropriate code at the beginning of the program.

In SPLIT_SUPPRESSION.PRG, the path and name of the source file, the format of the source file and a suppression file type value must be specified. The suppression file type determines the category of the suppression file that is being loaded. Please see Suppression File Types on page 77 for further information.

The example below shows how the code might look after it has been edited:

lcInputFile = "D:\Suppression_Load\MyFile"

lcInputFormat = "DELIMITED WITH CHARACTER '|'"

lcSID = "C"

In this example, "lcInputFormat" (representing the format of the source file) has been set to expect PIPE delimited.

In CREATE_FINAL_SUPPRESSION_FILES.PRG, it is only the path and file name that must be specified.

Suppression File Types

The suppression file type, specified by **lcSID** in SPLIT_SUPPRESSION.PRG links the file to a record in SUPPCATS.DBF, located in the matchIT root folder. This can ultimately be used to determine the information displayed on suppressIT reports next to that suppression file - the priority the suppression file takes amongst other suppression files and the hit costs that may be involved in using a specific file.

If a file is simply to be referred to as a 'Custom Suppression File' with no hit charges, then the file type should be set to "C". Otherwise, it should be set to correspond with a new record in SUPPCATS.

SUPPCATS.DBF already caters for common types of suppression files. The existing records can be used as examples of how to add further categories if required.

Running a Suppression Load Job Script

If the steps detailed in the "Before Running a Suppression Load Job Script" section have been followed correctly, the script will be ready to run.

To run a Suppression Load script, select Automation from the main matchIT screen. When the Job Maintenance dialog appears, select the job script that is to be run, from the "Job Reference" drop-down menu and click on "Start Job". The script will start. Unless the current suppression load directory has never been used to load suppression files before, the script will run without user intervention until the suppression load process is complete.

Note: when loading suppression files containing several million records, the scripts are expected to run for several hours before completion.

Running a Suppression Load Job Script for the First Time

If the directory containing the suppression source data has never been used for suppression loading before, the script will create a "Load" sub-folder and will copy into it the required intermediate load tables.

After preparing the new load directory, the script will automatically update its paths and file mapping to the new suppression load directory, and the following message will appear:

Upon clicking 'OK', in order to recompile the script with the new file mappings, matchIT will close. At this point, restart matchIT and re-run the job script. It will now run until all suppression files have been loaded.

Using the Final Suppression Files

When the script has finished running, the final suppression tables will be ready to use.

These can be used just like any other suppression file in the suppressIT module; simply map the suppression file(s) in the Advanced suppression options and either tick 'Enable Automatic Suppression During Import' when importing a new file, or select 'Suppression>Find Suppressions', to suppress a pre-imported file.

For further information on using the suppressIT module, please refer to the suppressIT section of the User Manual on page 60.

Matching

Introduction to Matching

In matchIT, there are two critical steps: the first step is the Import or *Extract* process, this process performs (optional) *data enhancement* and generates *Match Keys*. The second step is the actual *Find Matches* step itself. In turn, the Find Matches step can be broken down into two steps as follows:

Match Key – Definition

In an ideal world, a user would be able to compare every single record in a database with every other record in the database. In a 10,000 record database, this would mean you would do roughly $10,000*5,000 = 50,000,000$ comparisons – even on this small database, the processing time would be unreasonably long.

Another problem with comparing every single record with every other record is that you may find numerous false matches – records which are obviously not matches to the human eye, but where certain key elements are either the same or empty. For example, if the name and part of the address are the same, and the zip code is blank, then it may add up to give a matching score. Although this score may be fairly low, pairs of records like this should not usually be reported as a match at all, because they would obscure genuine ("true") matches with a lot of miskeying and inconsistency between the records: if too many false matches are reported, it becomes difficult for the user to "see the wood for the trees".

We need a way to improve on this process, both to cut down the sheer volume of comparisons, and also to reduce the number of false matches that may occur. The way to do that is by using *Match Keys* – a match key is something that groups of records in a database have in common, which indicates that detailed comparison of the records is worthwhile to see how well the records match each other in other respects. The records within these groups are only matching *candidates*, they may not be true matches – further inspection is necessary to check if they are true.

To reduce the number of comparisons, we can look at a field and say that only records that match the value of that field are potential matches. For example, we could select a Surname field (if the surname is held separately from the title and initials), this would only compare records of individuals with the same surname. This would instantly and drastically reduce the number of comparisons we would have to do. Now, for example, Mr Smith would not be compared to Mr Jones, because their surnames are completely different. All the Smiths would be compared, because there is the possibility that they might be the same person. In this approach, we have used the Surname field as a *Match Key*.

This approach is an improvement on our original idea, but it has its limitations. Mainly, it does not allow for miskeying of the surname. For example, take the name Shore. People generally expect this to be spelt Shaw, as this is far more common than Shore. The two spellings *sound* the same, but our solution above would not compare Mr Shaw and Mr Shore, because the surname field is different. The obvious solution to this is to use a "sounds like" version of the Surname field as a *Phonetic Match Key*. The *Extract* process takes important fields (such as name, address, company), and generates phonetic versions of the key elements in those fields.

However, even this approach has its drawbacks. The main problem is that this still gives us too many records to compare (most of the time). For example, if you are comparing records across

the whole country, there are a lot of Smiths – and Mr Smith in New York is obviously not the same as Mr Smith in San Francisco. Secondly, it does not cut down on the number of false matches. To get around this, we use combinations of fields to narrow the search. In the example above, it would be a good idea to consider part of the zip code code as well. So we may choose something like phonetic key of surname plus the outward bound (first half) of the zip code. This is a more explicit *Match Key* than just phonetic surname.

In matchIT, we do two or three scans of the database, using different Match Keys to find all the possible matches. If, for example, the two Mr Smiths above were the same, but one record did not have a zip code, we would miss the matching records if we relied just on phonetic key of surname plus the first half of the zip code. This is because the search requires the first half of the zip code as well as the phonetic surname to be the same, in order for the two records to be in the same "candidate group". However, if the street in both the records is the same (which will be the case if they are true matches), we can do a second search based on phonetic key of surname plus the phonetic key of the most significant word(s) in the street, this would find the match on the second scan of the database.

At the start, we mentioned that there are two steps to finding matches. The process we have described above is the process of establishing *Match Keys*. This helps us cut down the number of potential matches that are fed into the second step, which is the *Match Score* step – this step enables us to determine how well the other data matches in each pair of records.

Match Score – Definition

So, now we have two records that may or may not be for the same person (or family, household or company). If we have used the Match Key process above, then the likelihood of them being the same is much improved.

We will now go through the record, field by field, and work out how similar they are. Each field can contribute to a *match score*, depending on how close those fields in the records are. At the end, we have an overall score that tells us how alike two records are – the higher the score, the more similar the records are. When deleting duplicates with matchIT, you can enter a threshold score, for and above which matchIT will automatically flag one record from all matching pairs. With most data files, all pairs scoring above a particular score (say 90) will be true matches and anything below (say) 80 will be false matches – this leaves a grey area between (say) 80 and 90 where most of the pairs are true matches but some are false. For "underkill", you can therefore enter (say) 91 as a threshold score for deletion. For "overkill", you can enter 80.

This process works well, but it is essential that users fine tune it themselves. This is because matching requirements can vary from company to company, file to file and even job to job. For example, sometimes you want to dedupe to individual level, sometimes to company or family. In addition, data files vary widely in the structure and overall "shape" of their data". Sometimes zip codes are reliable and complete, sometimes they are unreliable. There are many ways that the data can differ or be incomplete, and this frequently changes with every new file that you "dedupe", or match against other files.

Because we know that everyone's data is different, we have allowed the way that two records are compared to be customized, using a parameter table that tells us how much each field contributes to the overall matching score. Using this table, we can tell matchIT how important each field is in the matching process. We call this the *Weights* table, as it reflects the relative weighting that each field has towards the total match score. For each field that the user wants to match on, there are five weights, representing whether the pair of fields in the two records is a *Sure*, *Likely* or *Possible* match, or whether *One* or *Both* of the fields are *Empty*. If a pair of fields is quite different, it is given zero weight i.e. it does not contribute at all to the total matching score.

For name and address matching, matchIT compares the elements of the name (or address) as a whole, rather than just comparing them element by element – this allows it to match names where

some of the components are omitted or in a different order in one record (e.g. John Michael Smith and Mike Smith, or addresses which have a house name in one record but not in another).

To summarize, the *Match Keys* act as a filter, filtering in candidate pairs that may be matches. Match Keys work best when using phonetic keys, and combinations of fields. After that, each candidate pair of records has a *Match Score* calculated for it, based on the *Weights* table. This allows matchIT to work out how well each pair of records matches.

Further detail on how matching works is given in the Online Help.

Summary of Finding Matches

To find matching records in a Main File, the following steps are used:

- Import the data. The simplest method is using the "Setup Wizard" on page 20. If you are not using the Setup Wizard, select Restore Standard Parameters from the dialog which you see when you select Import, and choose the appropriate matching configuration before you Import the data – see Importing without using the Setup Wizard (on page 24). If you are importing through the Setup Wizard, at the end of the Setup Wizard, tick the checkbox for "Create Match Keys" and choose the appropriate matching level in the "Select matching level" radial.

- Utilize the Find Matches option either at the end of the Setup Wizard, from the Matching menu, the toolbar, or from the dialog at the end of Import.

Find Matches button

- Choose the Match Keys to use (unless you chose Find Matches at the end of the Setup Wizard, in which case default Match Keys are used).

- Wait while matchIT searches for possible duplicate records.

- View, verify and/or flag the matches that have been found.

Matching Parameters

When importing data through the Setup Wizard, matchIT asks what level of matching is to be used. The choices offered depend upon whether the data is business or personal.

The parameters can also be modified for specific requirements. To change these, either select Save/Restore Setup (see "Multiple Parameter Sets" on page 16) or select "Options" from the Jobs/Setup main menu and change the "Matching Setup - see Online help for more information".

Matching Levels

The Setup Wizard allows you to choose between contact and business level matching, only if there is a field named "Company" in your Main File.

Residential Data

Individual Level

This will find matches at an individual level e.g. John Smith and Mary Smith living at the same address will not be matched, nor will John Smith and James Smith. However, John Smith and Mr J E Smith will be matched and (by default) John Smith and E J Smith will be regarded as a possible match, as perhaps E J Smith is known by his middle name.

Family Level

This finds matches on surname at the same address e.g. John and Mary Smith at the same address will be matched, as will John Smith and James Smith, and all but one record will be flagged.

Household Level

This matches records with the same address, regardless of surname e.g. John Smith and Lucy Jones living at the same address will be matched.

Business Data

Contact Level

At contact level, deduplication is performed down to one record per person at a location. This is effectively the same as individual level matching for residential data, as by default matchIT ignores the company name for contact level matching – this tends to work better because company names change so much, different companies within the same group often have employees in common etc. For example, John Smith and Mr J E Smith will be matched, even if one is at British Steel plc and the other at Corus, as long as the addresses and zip codes match well enough. However, John Smith and Fred Brown at the same address/zip code will not be matched

Business Level

This level is used to produce one record per company or business. Therefore, two different employees working for the same company will be matched, as long as the addresses and zip codes match well enough. Once you have selected to continue with Business Level matching, you will be prompted for Loose (the default), Tight or Legal business matching. These options are illustrated in the window prompt.

Address Level

This matches records with the same address, regardless of surname or business name. All contacts and businesses at the same address will be matched when using this level of deduplication.

Match Keys

Match Keys define which fields matchIT uses when looking for candidate matching pairs in your Main File – it is important that you understand what this means, so please refer to the definition of Match Keys in <u>Introduction to Matching</u> (on page 79) if you are unsure.

At the top of the screen is a scrollable list of Match Keys called the Fields/Key list. This contains both suggested Match Keys and all fields in the Main File, with the keys that you are most likely to need displayed at the top of the list.

Use the Basic or Advanced option to display these keys in plain English or as field names from the Main File. In their plain English description, keys detailed in proper case have been standardized by matchIT, ones entirely in lower case have not been standardized. It is better to use the standardized fields, but you could construct your own keys from non-standardized fields, in addition to using matchIT's generated keys.

The simplest way to select Match Keys is to check the "Use default keys?" box on the Match Keys screen. This will select a standard set of keys which are suitable for most jobs – as supplied on initial installation, they are suitable for matching at individual, family, contact or business level. The keys are suitable for data that is either virtually all US or is all non-US. For household matching, the default keys will work okay but are not ideal, as two of the supplied match keys feature phonetic surname and only one does not. Therefore, for household matching we recommend that you select from the first few keys in the List of Fields/Keys Available as follows:

- two keys that do not feature phonetic surname e.g. for US data:

  Zip code

  Address Key

  or (for non-US data):

  Address Key

  First 5 characters of ZIP code + Phonetic Key of Street.

- one key that does feature surname to pick up extra household matches where the surname is the same but for some reason the household match is too fuzzy to be picked up by either of the first two keys e.g. for US data:

  Phonetic Surname Key + First 5 characters of ZIP code

  or (for non-US data):

  Phonetic Surname Key + Phonetic Town/City Key.

The set of default keys provided with matchIT may be overwritten by checking the box "Save as default keys?", but don't do this unless you have a thorough understanding of how match keys work.

Selecting Specific Keys

To select your own keys:

- Click "Clear Keys" to clear all currently chosen keys, if required.

- Highlight the field you wish to use in the Field/Key list and click "Select Key" or double click and this field will be added to the key list.

- To add another field into this key, highlight the second field to be used and click "Select Key" again – in this way, you can build compound keys such as Zip code + Phonetic Surname Key.

- Click "New Key" to add another and repeat from step 2.

If you make a mistake, just click on the key that is wrong (in the "Keys Chosen For Matching" box) and click "Remove Key".

"Verify Keys" will check your chosen keys for you and suggest any problems which may occur with using those keys.

New Analysis

If you are working with a Main File that has already been analyzed for matches, you may not wish to lose any previous matching results. If so, uncheck the box "Is this a new analysis?"

Matching Summary

Matching information is recorded in the Perform (short for "performance") database. The Matching or Overlap Summary shows information from this database, which you see automatically at the end of the analysis, or you can view it from the Matching menu. The information displayed is as follows:

Database in use: the database you are using.

Number of records: the total number of records in your database.

Number of matching pairs: the number of pairs of potentially duplicate records that matchIT has found.

Number of potential deletions: the number of records matchIT will flag, provided you accept the results and flag at the default threshold score. This is often different to the Number Of Matches Found. This is because if you have more than two records which match each other, say 1, 2 and 3, then record 1 matches record 2 and record 1 matches record 3, but also record 2 matches record 3. This gives 3 matching pairs, but only 2 records will be flagged. This exponentially increases for larger match sets: 10 records in a match set means 45 matching pairs, but only 9 potential deletions!

Matches found by match key

Run number: When matchIT does its Find Matches routine it makes several passes (or runs) over the database, on each of the chosen Match Keys.

Records compared: the number of records matchIT has read for each match key. This may be less than the number of records in the Main File if there are records with blank keys, or a Start Range was specified.

Matches found: the number of matching pairs found by each key.

Match Keys used: the Match Keys used since the New Analysis box was last set On.

Matches found by score range

Score range: This shows ranges of match scores depending on the *Minimum Score To Report* and the maximum possible score.

Number of matches in range: This is the number of pairs of matches found in each score range.

View/Verify Matches

After matchIT has performed the Find Matches step, the following screen is displayed:

From here, the matches found can be viewed, verified or flagged. These options can also be accessed from the Matching menu. View Matches is used to simply produce a report listing the matches found whereas Verify Matches allows matches to be flagged interactively, declared false matches, and/or allows records to be modified. Flag Matches removes all duplicates scoring above a threshold match score, without further user intervention.

View Matches Window

Selecting View Matches displays another window, as shown:
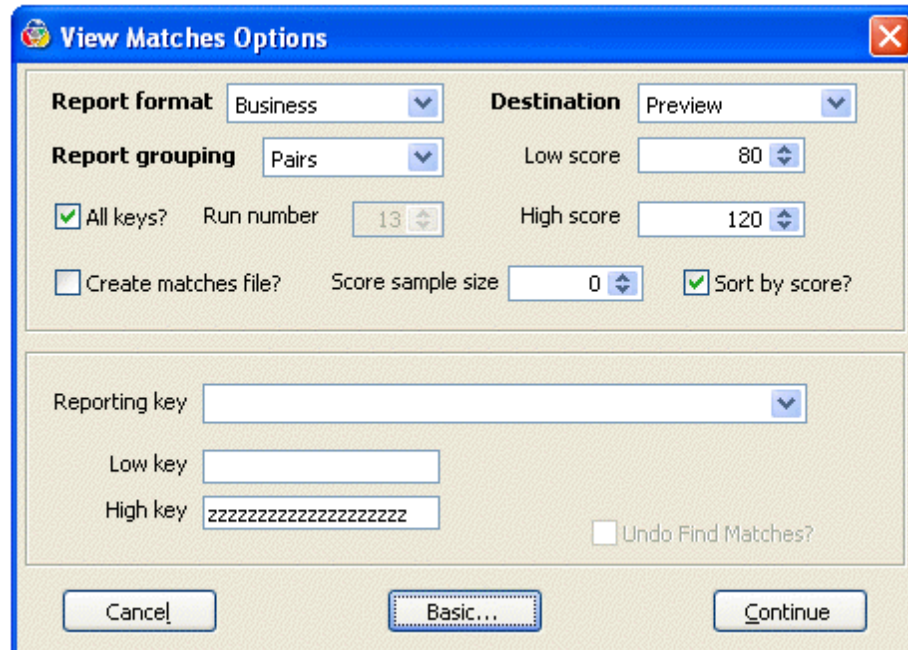
The most useful options are those in the top half of this screen:

There is one run for each matching key selected.

- The results from all runs (since the New Analysis box was checked) can be viewed, or one run number can be selected by unchecking the "All runs?" box. The default run number is that of the last matching run performed.

- Matching records can be grouped in pairs (e.g. John Smith, Mr J Smith and J Smith Esq at the same address will be shown as three pairs of matches), or in sets where all records matching each other are displayed together i.e. the three Mr Smiths above would be shown as one set of three records.

Potential duplicates are given a score based on how well they match: the higher the score, the closer the match

- The range of matching scores for which the results should be displayed is also controllable and (if viewing in Pairs) records can be sorted by score. Set the score sample size to e.g. 10 if you wish to see just the first 10 pairs for each match score. Selecting this automatically sorts the records by score and displays them in pairs

- The report format will default to Business or Residential, dependent upon whether there is a company field in the Main File. (See "Matching") You can choose your own report layout from here and also Verify Matches (see next section).

If you want to print the report to a printer other than the default printer, or to print more than one copy or a page range, you should select a Destination of Printer, rather than print from the Print Preview.

- Destination can be either:

  - Preview for a print preview,

  - Printer to print the report,

  - File output to a text file or

  - PDF to save a PDF version of the report.

Verify Matches Window

Verify Matches provides an easy way of checking the quality of the matches found, even if you want to flag matches globally. Selecting **Verify Matches** from the **Matching** menu or other dialog shows all matches in order of score, lowest score first. Verify Matches is also available as a "Report Format" from the View Matches (on page 86) screen, where you can specify which runs to view, the score range and score sample size.

The Verify Matches option within matchIT has the ability to display matching records in either a pairs view (default) or a sets view. To change the default view for the Verify Matches option you will need to relabel a file named "NotVerifySetsFromWizard.txt" in the matchIT directory (C:\Program Files\matchITv51). Essentially the file should be named "NotVerifySetsFromWizard.txt" if you wish to examine records in the pairs view. However, if you wish to look at records in a sets view the file needs to be renamed to "VerifySetsFromWizard.txt".

For global deletion, see "Flag Matches" on page 92. The default screen used for pairs of matches is shown below:

Verify Matches in Pairs
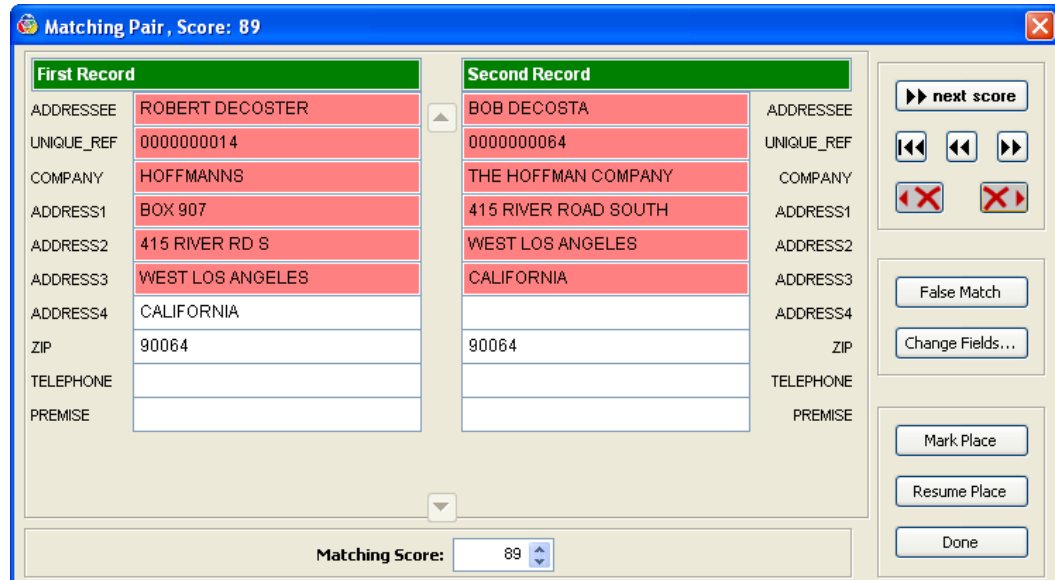
This screen will display potential duplicates, in pairs. The different colored highlighting is used to help distinguish where the differing fields are in the records. Fields in red are different, yellow shows information which is contained in the same field in the other record and white displays those fields which are identical.

The buttons in the top right part of the screen control movement through the table.

         goes back to the first pair displayed (lowest matching score selected).

         goes back to the previous pair displayed.

Keyboard shortcut keys for next pair and previous pair are Page Up and Page Down.

         goes forward to the next pair (which may or may not have the same score as the current pair displayed).

         jumps to the next pair with a different match score. If you are viewing the records in order of score (the default), this will be the next highest score. (A little message "No more scores!" appears in the top right when the end of the pairs is reached.)

Below this pair the matching score is shown and (on the right), various buttons for dealing with a matched pair:

Keyboard shortcut keys for Flag/Recall Left and Flag/Recall Right are Ctrl+left arrow and Ctrl+right arrow	◀✖ ✖▶	To flag a record in Pairs format, select one of these buttons to flag the record on the left or right-hand side respectively. The button then changes: to recall a record, select the appropriate button with a green tick.
False Match is Ctrl+Delete	False Match	If the pair shown is not a true match, you can select this button to remove the match – meaning that this pair will no longer be regarded as a match.
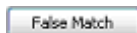
Typically, you either flag false matches in a score band if most of them are true matches, or flag the odd true match if most of them are false. You do not need to both flag false matches and flag true matches within a score band, as global deletion will flag all matches above a match score as long as they have not been declared false. If you are using the Matrix Report or Group Matches feature, any matches you flag interactively will not be included. To allow this, you can increase the match score of any pair in the grey area, so that you move them into the area that is being automatically flagged or grouped.

	Change Fields...	This button allows you to enter the name or names of any additional fields that you want to see in the Verify window. Just scroll down and type the name of the field at the end of the list. You have to know the exact spelling of the field name or it will not be displayed, but you don't have to specify the field type or width. You can also change the order in which the fields are displayed, by dragging the square button to the left of the name. Then select OK and say 'Yes' to the question "Make structure changes permanent". NB: this question refers to the structure of a temporary work file, not the structure of the Main File.
	Mark Place	This button asks matchIT to remember the current pair so you can return to this pair later if you are deleting a lot of matches interactively.
	Resume Place	This button will return you to the marked pair.

Click "**Done**" when you have finished. If you have not marked your position in the file, matchIT will ask if you wish to. This means you can return to the same point at a later date.

Verify Matches in Sets
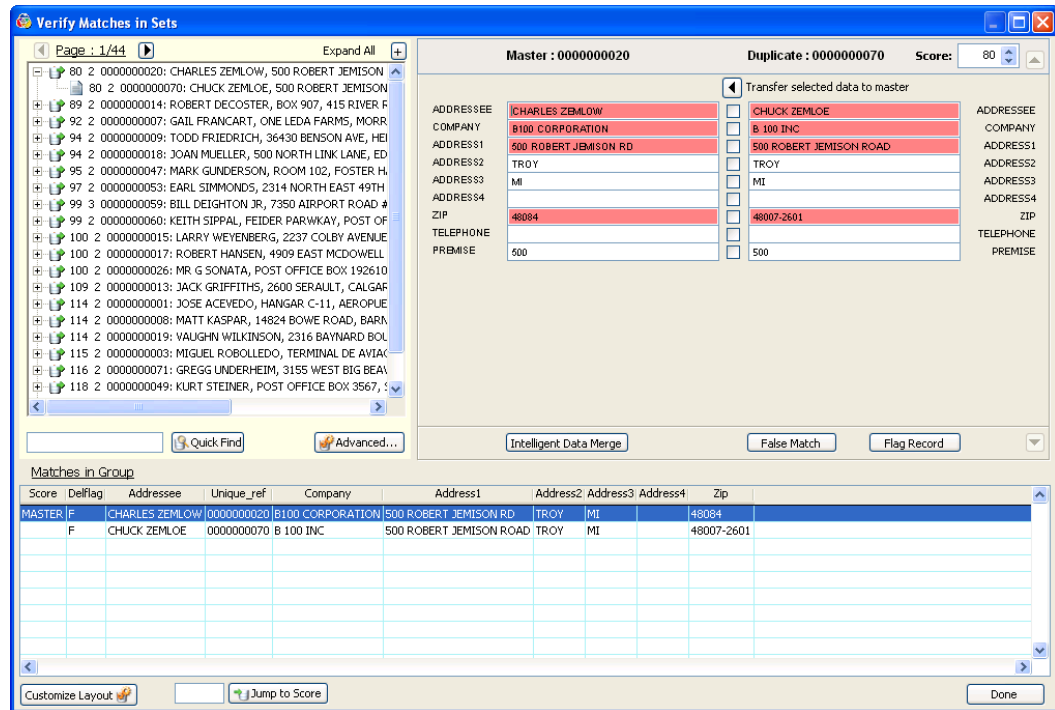
This screen will display potential duplicates, in sets. The different colored highlighting is used to help distinguish where the differing fields are in the records. Fields in red are different, yellow shows information which is contained in the same field in the other record and white displays those fields which are identical.

The tick boxes between matching records (in the top right part of the screen) control what information is transferred to the Master record.

<table>
<tr><td>☐</td><td>specifies what fields to transfer to the Master record.</td></tr>
<tr><td>◀</td><td>transfers selected fields to the Master record.</td></tr>
<tr><td>◀ Page : 1/44 ▶</td><td>goes forward to the next page of matching sets.</td></tr>
<tr><td>Expand All +</td><td>expands all matching sets shown in the current page of matching sets.</td></tr>
</table>

Below the matching pair display there are various buttons for dealing with that particular matched pair:
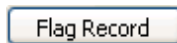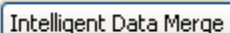
<table>
<tr><td>Flag Record</td><td>to flag the duplicate record in the pairs view section, select this button. The button then changes: to "Restore Record" which allows you to unflag the record.</td></tr>
<tr><td><i>False Match is Ctrl+Delete</i></td><td>False Match</td><td>if the pair shown is not a true match, you can select this button to remove the match – meaning that this pair will no longer be</td></tr>
</table>
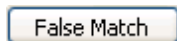
90
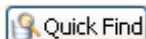
| | regarded as a match. The button then changes: to "Restore Match" which allows you to reactivate the displayed matching pair. |
| Intelligent Data Merge | uses the Intelligent Data Merge (see Intelligent Data Merge) settings specified by the user to create a meta-record of the two records shown in the pairs section. |
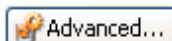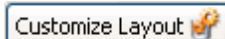
Typically, you either flag false matches in a score band if most of them are true matches, or flag the odd true match if most of them are false. You do not need to both flag false matches and flag true matches within a score band, as a global deletion will flag all matches above a match score as long as they have not been declared false. If you are using the Matrix Report or Group Matches feature, any matches you flag interactively will not be included. To allow this, you can increase the match score of any pair in the grey area, so that you move them into the area that is being automatically flagged or grouped.

Quick Find	quickly find records that contain a piece of information as specified by you. The records that contain the information will then be displayed in the Verify Matches in Sets window.
Advanced...	customizes the "Verify Matches in Sets" window by allowing the user to apply filters and sort orders so that only records of specified concern are displayed.
Customize Layout	allows you to enter the name or names of any additional fields that you want to see in the "Matches in Group" section. Just scroll down and type the name of the field at the end of the list. You have to know the exact spelling of the field name or it will not be displayed, but you don't have to specify the field type or width. You can also change the order in which the fields are displayed, by dragging the square button to the left of the name. Then select OK and say "Yes" to the question "Make structure changes permanent". NB: this question refers to the structure of a temporary work file, not the structure of the Main File.
Jump to Score	jumps to the first matching set of a specified match score.
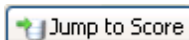
Click "**Done**" when you have finished.

Intelligent Data Merge

Introduction

The Intelligent Data Merge allows a user to construct a list of simple rules for the merging of matched records. These rules can then be applied to the merging of fields as part of other operations performed on the matching records. The idea behind the process is that it allows the user to extract the most important information from each pair of fields and combine that information into one field of superior quality. These enhanced fields are then inserted into the appropriate record (i.e. the destination record or record to be kept), thus improving the overall quality of the record.

Intelligent Data Merge configuration should be reviewed (via the Jobs/Setup menu, Matching Setup) if you change matching parameters, to avoid incorrect merging of data. If in doubt, you should not use this option. Thorough checking of results is advisable when using Intelligent Data Merge, especially on any new type of data.

Functionality

The process compliments the functionality of the *Transfer Data*, *Flag Matches* and *Verify Matches* operations by giving the user more control over what data should be retained.

Transfer Data

The Transfer Data operation (from the Matching or Merge/Purge menus) allows a user to transfer data from specified fields in one record to those fields in the matched record. The IDM process can override the options for the Transfer Data operation. To use IDM in the Transfer Data option, simply check the appropriate box on the main Transfer Data option screen.

Flag Matches

The deletion process employs a Deletion Priorities table to select which record should be flagged. The IDM process will not override this, rather complement it by running after the choice of which record to flag is made, thus retaining valuable information in the flagged record and merging it with the kept record. To use IDM when deleting matches simply check the appropriate box on the screen that appears at the beginning of the deletion process.

Verify Matches

Within the Verify Matches operation, the IDM process operates as a tool which the user can control interactively.

Selecting this option will produce a "meta" record showing the result of the process on the selected fields. The process should take no more than a second or two and then the window on the right will appear. The window shows the results of the intelligent merge for the two records. Only the fields which exist in both records and have been specified in the IDM Options set-up will appear. Clicking either of the '**replace**' buttons will insert these fields into the corresponding record and additional changes can then be made. Selecting **Discard** will return to the Verify view without making changes.

The Options Screen

The main options screen for the IDM process can be found under the menu **Jobs/Setup** > **Matching Setup as Intelligent Merge Options**. This is described fully in the Online Help.

Additional Information

Certain fields, such as Address and Telephone, are dealt with in a different way to others in the merging process. The overall way in which these fields are treated cannot be altered. Please see the Online help for further information.

Flag Matches

Flag matches is used to flag records above a certain matching score without further user intervention. It can be selected from the screen displayed after finding matches or from the Matching menu.

If selected from the main menu, you are first asked if you wish to modify any of the scores (e.g. to flag false matches where you know the reference numbers of the pairs involved). Also, you can view how many pairs achieved each match score and flag as False any particular match score – for details on this, see
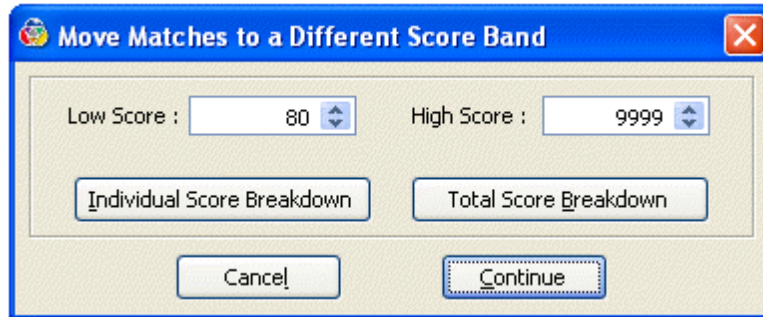
Modifying scores is not usually necessary: click Next if you do not wish to change anything.

The "Flag Records" screen is displayed next (or straight away if you chose this option after finding matches). Here you enter the matching score at and above which you want duplicate records removed. It defaults to the
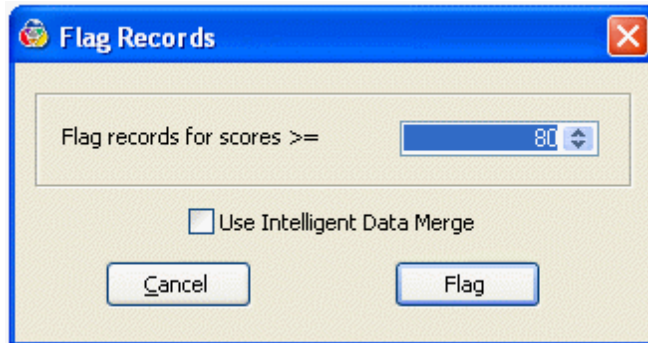
If you leave the score at the default, only those records that weren't given a Matching Score will be left in the Main File. If you enter e.g.100, it will only flag records in the Main File with a Matching Score greater than or equal to 100, leaving one record for each individual, company, family or household, depending on the level of matching you selected previously. You can also select to use

Transfer Data

The Transfer Data operation (from the Matching menu) allows a user to transfer data from specified fields in one record to those fields in the matched record. However, the IDM process can override the options for the Transfer Data operation. To use IDM in the Transfer Data option, simply check the appropriate box on the main Transfer Data option screen.

Below the field picker there are other options can be set:

- **Transfer for scores** – data is only transferred for records whose matching score is equal to or greater than the value shown.

- **Transfer data** – from Higher to Lower shows where the data will be copied from to. Clicking on one of the buttons (Higher/Lower) will swap the order.

- **Always Replace/Replace Only if Empty** – If there is data in a field that is about to be overwritten, matchIT needs to know whether to replace it (always replace) or leave it (replace only if empty).

- **Blank All Fields/Append to Existing Fields** – You can optionally choose to blank all the fields you are transferring to before starting, or append the data being transferred to that already in the field.

- In all cases, you should ensure that the field being transferred to is large enough to contain all the data or some of it will be lost. To make sure that only the best match is

used to transfer data across where many records match one, define a field MATCHSCORE (numeric 8) in your receiving file, using "Main File Layout" on page 30.

- Click "Continue" to start the process.

Group Matches

Group Matches is available from the Matching menu, and is used to collate match sets after a matching run. For this to happen there must be a MATCH_REF field in the Main File layout, which is of type character and is at least as wide as the UNIQUE_REF field. matchIT will ask for a threshold score, and then collate all records which are in a match set with a matching score higher than it. It does this by determining the 'master' record of a set; that is, the one which is to be kept in favor of all the others in the set. It then copies the contents of the master record's UNIQUE_REF field into the MATCH_REF field of all of the other records in the set. This means that if the Main File is viewed in MATCH_REF order, all the match sets will be grouped together.

matchIT will do the exactly the same thing as described above if 'Flag Matches', is chosen from the Matching menu, but will also flag the matches. Group Matches will **not** flag any records.

Delete Suppression Records

If you have processed a job created by the Multiple File Wizard, you are prompted at the end to remove suppression records from the mailing file. If you don't say "Yes" to this prompt then, you can remove them via this menu option.

Total Score Breakdown

Total Score Breakdown is available from the Matching or Merge/Purge menus, or before you select Delete Matches or any of the Merge/Purge options from the relevant menu. This facility shows the matches found on the currently selected Main File (or Files) categorized into score bands. It displays the number of matches found at a particular score, an example of a matching key that was used to find them, and a flag value, which can either be True (T) or False (F). Setting this flag to **False** marks as False matches all of the matches found at that score, so they will not appear on any future matching reports. This is useful when, after verifying or viewing the matches, it is found that matches at a particular score are all False, but some lower scores are True matches.

Matching Setup

This section is for users who wish to customize matchIT's matching setup. We recommend that you have some experience with matchIT before changing the default setup.

Basic Matching Options

Must have match on gender

If you set this parameter "on", then matches where the sex is different (where the sex is normally deduced from the title) will be suppressed, even if the name matches and there is no independent confirmation of the sex by the forename. Note: even if you have this parameter off, matchIT will always return a nil score on name (if you are not doing family matching) if the sex of the two records is different and is independently confirmed by the sex deduced from the forename of each record.

Sex is only regarded as different if one record must be Male e.g. Mr and the other Female e.g. Mrs, so it will match Mr and Dr, or Mrs and Dr etc. but not Mr and Mrs.

Must have match on location

If this parameter is set "on", then a matching pair is not reported if a 'location check' has failed. In detail, this means that the zip codes in the two records (if present) must achieve at least a probable match with the address score at least a Possible match, or the address score must be at least a Likely match irrespective of the zip codes, or the zip codes must achieve a Sure match irrespective

of the address. This is to prevent false matches where there is some match on address, but where the addresses are clearly not the same, for example "10 High Street, Bookham", and "10 High Street, Alford". Switch this parameter **off** if you want to match people or companies in different locations; you may want to match on items of data that are independent of location, such as date of birth or bank account. See also the section about Address Matching.

Must have building/house/apartment number match

If you set this parameter "on", then matches with different premise or apartment numbers in the addresses will be suppressed. This could be appropriate if you have a very localized file with a lot of neighboring records, perhaps with similar names.

Don't allow matches where there is no street number in one record

If this parameter is "on", then matches where one of the records has a premise number in the address, and the other hasn't will be suppressed. This is useful when your file has a lot of addresses with house names in the address lines. Note that this parameter has no effect if the "Must have premise match" parameter is off.

Flag higher reference

(does not apply to matchIT **Lite**)

This specifies the default rule for global flagging of matches, when the deletion priorities are equal or switched off (see [Keep best record](#) and Online help). If you set the Flag Higher Reference parameter "on", the record with the higher reference number from each pair is flagged by the Flag Matches option; otherwise the earlier record in the file is flagged.

Name key generation

This parameter governs how the NAME1, NAME2 and NAME3 Match Keys are generated; if the parameter is set to 'Personal', the keys are generated from the personal name, if it is set to 'Business', they are generated from the company name (See the '[Phonetic name matching](#)' parameter on page 100). If you want to switch from any other type of matching to business matching or vice versa, you should do so by using the Save/Restore Setup (see Online help for more information on Save/Restore Setup) option, rather than by just changing this parameter. This is because business matching uses different matching rules to other matching levels.

Minimum score to report

The Minimum Score to Report is the minimum matching score necessary to allow the records to be reported as potential matches. These records are written to the Matches table (or the Merges table for Find Overlap processing).

Phonetic name matching

There are two stages to the matching process that matchIT uses; the key stage and the scoring stage. The first stage creates standardized and phonetic keys based on the input data, which allows potential matches to be identified. The second stage scores each pair of potential matches, using phonetic and fuzzy matching. This parameter governs the phonetic algorithm that matchIT uses when generating keys and for scoring.

There are three choices available:

- **soundIT**

  matchIT provides a unique phonetic algorithm for name matching, called *soundIT*. *soundIT* takes account of vowel sounds and syllables in the name, and, more importantly, determines the stressed syllable in the word. This means that "Batten" and "Batton" sound the same according to *soundIT*, as the different letters fall in the unstressed syllable, whilst "Batton" and "Button" sound different, as it is the stressed syllable which differs. Another advantage of *soundIT* is that it can recognize groups of vowels and consonants that form vowel sounds – thus it can equate "Shaw" and "Shore", "Wight" and "White", "Naughton" and "Norton", and "Leighton" and "Layton" (which are all reasonably common English surnames).

  This algorithm was developed with extensive testing on a large table of the most common surnames in the UK. Therefore, it is specifically designed to be used with English names. If a file with mostly non-English names is processed through matchIT, then you may want to try the 'Loose' *soundIT* or Soundex algorithms instead. For US data we recommend that you use *soundIT*, because it is proven to work well also with Spanish, German and other names that occur commonly in the US. *soundIT* has been designed with foreign language versions in mind (i.e. for data collected in countries where foreign languages are spoken). These could quite easily be developed, according to demand. Please contact your supplier if you are interested in this.

  Note that the keys that matchIT generates are 'Loose' *soundIT* keys, where all vowel sounds are equated, together with some consonants, such as 'm' and 'n', 'd' and 't', 's' and 'f'. This is so that potential matches are not missed at the key stage; matchIT uses the 'full' *soundIT* algorithm at the scoring stage, which will separate out false matches from true matches.

- **Loose soundIT**

  This option is effectively the same as the *soundIT* option, except that matchIT uses the 'Loose' *soundIT* algorithm as described above at the scoring stage. This is for use mainly with non-English names, on which *soundIT* works less effectively, and can miss true matches. This option should not be used on files with mainly English names, as it can potentially lead to more false matches.

- **Soundex**

  Soundex is a widely-used algorithm (patented just after the First World War!), which constructs a crude non-phonetic key by keeping the initial letter of the name, then removing all vowels, plus the letters H, W and Y, and translating the remaining letters to numbers. It gives the same number to letters that can be confused e.g. 'm' and 'n' both become 5. It also drops repeated consonants and consecutive letters that give the same number e.g. S and C. It only takes the first four characters of the result, or pads it out with zeroes if it is less than four long. Thus all the common spellings and misspellings of the name "Tootill" equate to the same Soundex key: Tootill, Toothill, Tootil, Tootal, Tootle, Tuthill, Totill are all translated to "T340".

  The algorithm that matchIT uses is an enhanced version of Soundex, and is for use mainly with non-English names. This option should not be used on files with mainly English names, as it can lead to false matches e.g. Brady, Beard and Broad get the same Soundex key.

- **Non-phonetic**

  This option constructs a non-phonetic version of the supplied name fields as match keys and allows only non-phonetic name matching.

Note that, at the scoring stage, matchIT performs name comparisons using data from the NAME field, not from the phonetic keys NAME1, NAME2 and NAME3 – this way it can check for simple typing errors such as "Wilson" and "Wislon" which do not match phonetically.

Keep best record

This parameter governs the use of deletion priorities (see Online help), where matchIT makes a decision about which is the 'best' record to keep from a matching pair e.g. the one with the most data. There are three settings: Disabled, Internal Only, and Internal and Overlap. Set the parameter to Disabled if you don't want to use deletion priorities, and want to use default rules (see the 'Flag Higher Reference' parameter on page 98). Otherwise set it to Internal Only, where deletion priorities will be used on matches in a single file, or Internal and Overlap, where they are used on matches in a single file and across files.

Advanced Matching Options
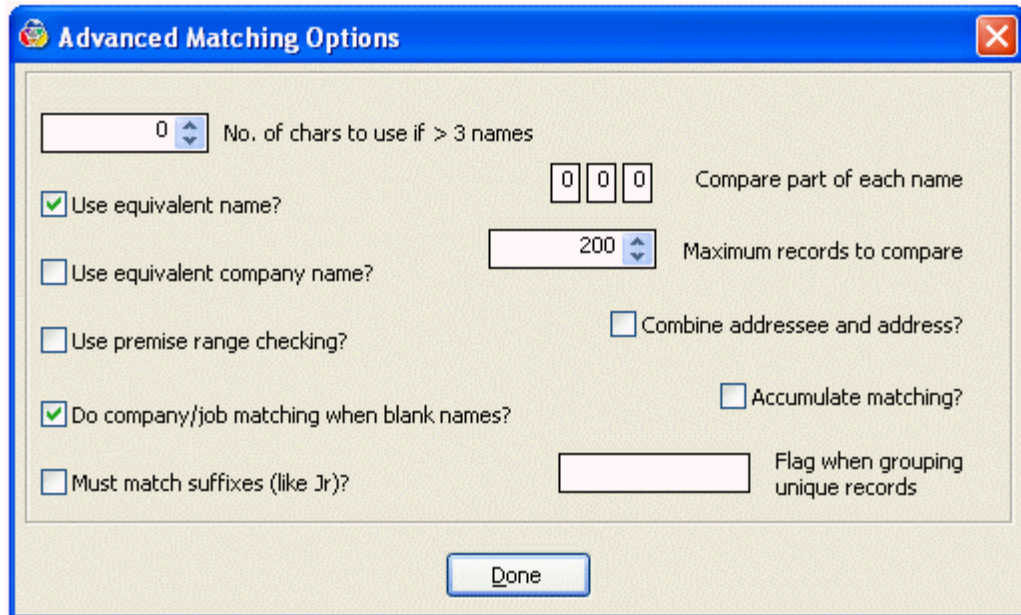
The Advanced Matching Options are described in the Online Help.

Modifying the Weights Used for Matching

Select the Weights option from the Jobs/Setup; Matching Setup menu if you wish to alter the values of the weights attached to the different fields in the matching process. You should not modify the matching weights unless you have read and understood the description in "Introduction to Matching" (on page 79).

A full description of Matching Weights and the Matching Matrices is given in the Online Help.

Deletion Priorities

The Flag Matches option deletes the records with the lowest priorities from each set of matches. If priorities are equal, it keeps the latest record on the file from each set or the one with the lowest Unique Reference, depending on the value of the Flag Higher Reference? Basic Matching option.

A full description of Deletion Priorities is given in the Online Help.

- **Field Name** is the name of the field in the selected Main File - if any field is not present, the rule for that field is ignored.

- **Comment** is just for your guidance, to explain if necessary what the field name contains.

- **Priority** shows the value which is accumulated for each rule that applies to the record being evaluated. When deciding which record to delete, matchIT accumulates priorities for each record and keeps the record with the highest total priority.

- **Value** specifies what Value the Field Name must be to obtain the Priority shown. The usual Value is "Empty" e.g. in the first rule above, if the surname (or first business name when matching business names) is empty, a Priority of -99 is given, which downgrades that record heavily. The other common Value is a constant e.g. if you specify rules as follows:

- There are some special Values which only apply to specific field names:

  - "Default" applies only to the salutation field: the record will be downgraded in the example above if the salutation is the current default e.g. Dear Customer rather than e.g. Mr Smith.

  - "First Name" and "Initial" apply to the first forename field, where it is a full name or an initial rather than empty. "First Name" can also apply to the 2nd significant word in a business name, when matching business names.

  - "Derived" applies only to the Prefix field: if the Prefix is derived from the first forename e.g. Ms for Christine, the record will be downgraded compared with a supplied Prefix such as Mrs or Dr.

If you define a field called DEL_PRI (numeric), matchIT will record the priority of each record that it examines for potential deletion, so you can check that your priorities are working as you expect.

Multi Level Matching

Multi–Level Matching is performed using a job script (see Job Scripts on page 114 for more information), which incorporates a number of bespoke Visual FoxPro programs, as well as customized settings and modified match keys.

The script is supplied as a template, and by default is set up to use 'Example1.TXT' and a corresponding DBF file. When the script needs to be used for processing live data, the user will need to modify the script in such a way that it points to the source data and corresponding DBF file(s). Once the user has pointed to their files, the script will be ready to run.

The job script will perform an Import and dedupe, it will then modify the structure of the working DBF and will perform the deletion step a total of three times. Upon finishing, the working DBF will contain separate fields displaying individual level, business level and address level matches. These field will enable the user to see which records match which, at what level, and how many other records any one record may match.

When the script has finished, the user can either export the data and matching results manually, using matchIT's standard Output menu, or can set up another script specifically for Output. This output script can be set up to run as soon as the Multi–Level Matching script finishes.

Breakdown of Job Script Rows

Each row displayed in the Job Editor, when browsing the Multi–Level Matching job script, is listed and described below. The Job Editor dialog can be accessed by selecting the job script from within the 'Job Maintenance' dialog, then clicking on 'Modify'.

Most rows in the script feature a program, and as these programs have been designed not to require any further modification from the user, the descriptions below merely give an overview of what each program does. For further information, please contact helpIT systems (see page 173).

- **Row 1 ("1 RESTORE SETTINGS.PRG")**

  Restores all parameters, weights and matching matrices required to find matches on all levels: NAME, COMPANY and ADDRESS. The parameters will force matchIT to generate name matching keys using individual names. Matching does not take place at this stage.

- **Row 2**

  Imports the source data file into a pre–defined DBF structure and generates match keys. At this step, the options (accessed via the 'Options' cell) must specify the correct input format.

- **Row 3 ("2 RENAME NAME FIELDS.PRG")**

  Renames the name matching keys (generated during the previous step using individual names) from NAME to IND_NAME, NAME1 to IND_NAME1 and from NAME2 to IND_NAME2. New blank Name fields are generated. The parameter that specifies which type of names are used to populate the name key fields are changed from 'Individual' to 'Business' in preparation for the next step.

- **Row 4**

  Regenerates name matching keys following the parameter change in the previous step. Keys are generated using business names.

- **Row 5 ("3 ADD IND WEIGHT.PRG")**

  Modifies the weights table, adding a weight on individual name as well as business, address and zip code.

- **Row 6**

  Finds matches using specific match keys, which incorporate phonetic company name, phonetic individual name, postcode elements and phonetic address keys.

- **Row 7 ("4 COPY MATCHING RESULTS.PRG")**

  Finds matches at address level using an additional program, which re–evaluates address–matching using different matching options from the previous matching run.

  With all levels of matching now performed, the program in turn deletes and recalls matches at each level, each time populating the MATCH_REF and SET_DUPS fields with matching information that corresponds to that level. After each Delete step, the MATCH_REF and SET_DUPS fields are renamed accordingly; after deleting individual level matches, these fields are renamed 'INDIV_REF' and 'INDIV_DUPS', after business level deletion; 'COY_REF' and 'COY_DUPS' and after address level deletion; 'ADDR_REF' and 'ADDR_DUPS'.

On completion of this program, the Multi–Level Matching script finishes. The processed table now contains references grouping matches together at the level they matched at.

Before Running the Multi–Level Matching Script

Before the Multi–Level Matching Script is run, it must be set up to point to the file that is to be processed. Row 2 of the script, which by default points to Example1.txt in the 'Source File' column, should instead point to the correct source data file.

All rows that point to Example1.dbf in the 'Main File' column, should point to the template DBF file corresponding to the source data file. This file will have the source data file imported into it and will then undergo processing, so it's structure must therefore match the source data file.

The easiest way to set up a template DBF file is to bring the source data file through the Single File Wizard. After mapping all fields as required, tick 'Create Match Keys' in the Processing Options dialog, and then click on 'More' and 'Save and Exit'.

If the Multi–Level Matching script is to be used more than once on a set of data, it is recommended that the script be cloned first. This way, after processing the data, the script will not need to be changed, and potentially changed back the next time that set of data needs processing. Scripts can be cloned by right clicking on the grid in the Job Editor dialog and choosing 'Clone Job'. The clone should be given a name that relates to the data that it will be processing.

Running the Multi–Level Matching Script

Once the script has been set up correctly, it can be initialized by clicking on 'Start Job' in the Job Maintenance dialog. The script should only ever be started from the beginning, as opposed to resuming it from a specific step using 'Resume Job', which will undoubtedly cause it to fail.

To run the script from the command line, please refer to 'Running a Job Script automatically when you start matchIT' on page 124.

Layout of Processed Data

When the script has finished running, the processed DBF file will contain new fields grouping the individual, business and address level matches together. These are structured in the same way as MATCH_REF and SET_DUPS.

If a record matches one or more other records at a specific level, the corresponding …DUPS field will display a figure representing the number of records that the record matched against. If it does not match any records at that level, the field will be empty. E.g. If record A matches records B, C and D at address level, it will contain a number 3 in its ADDR_DUPS field. If it only matches B and C at business level it will contain a number 2 in the COY_DUPS field. If it does not match any of these records at individual level, it will not contain anything in its INDIV_DUPS field.

The …REF fields indicate which records match which. In the above example, records A, B, C and D would have identical ADDR_REF values. A, B and C would have identical COY_REF values (record A would have its COY_REF field populated with its UNIQUE_REF value) and all of these records would have different INDIV_REF values (all identical to their UNIQUE_REF values).

The image below illustrates this example:

Addressee	Company	Address1	Address2	Address3	Address4	Postcode	Indiv_ref	Indiv_dups	Coy_ref	Coy_dups	Addr_ref	Addr_dups
JOHN SMITH	STANTON COPSHAW DIRECT	3RD FLOOR	LOW PROFILE HOUSE	12 NEWHALL STREET	BIRMINGHAM	B3 3ER	0000000001	0	0000000001	2	0000000001	3
MR F BLOGGS	STANTON COPSHAW DIRECT LTD	LOW PROFILE HOUSE	12 NEWHALL STREET	BIRMINGHAM		B3 3ER	0000000002	0	0000000001	2	0000000001	3
SARAH JONES	STANTON COPSHAW LTD	12 NEWHALL STREET	BIRMINGHAM			B3 3ER	0000000003	0	0000000001	2	0000000001	3
MISS A GREEN	FRINTON WILKES PLC	12 NEWHALL STREET	BIRMINGHAM	WEST MIDLANDS		B3 3ER	0000000004	0	0000000004	0	0000000001	3

Further Processing

Upon completion of the Multi–Level Matching script, the processed data can either be exported manually using matchIT's standard Output functions or it can be exported using another job script, which will need to incorporate the program 'OUTPUT.PRG'.

To manually export the data when the script finishes; return to matchIT's main menu. Make sure that the processed table is open in matchIT, then select Output>Output To File.

To automatically export the processed data using a job script, the additional job script will need to be set up before running the Multi–Level Matching script. The user will need to point to 'OUTPUT.PRG' from within the script and need to edit it to ensure that the correct output format, fields and destination(s) are specified. After setting the Output script up, open the Multi–Level Matching script in the Job Editor and tick the 'Run Another Job After This One' box, in the bottom right–hand corner of the screen. When prompted, select the Output script. Now, when the Multi–Level Matching script finishes the Output script will automatically run.

Maintenance

Adding A Custom Weight

If custom weights need to be added, the "3 ADD IND WEIGHT.PRG" program will need editing. The code required to add a custom weight is already used in the program to add a weight on Individual Name. Therefore, this existing code can be copied and pasted, and then edited to look at the field in which the weight is to be added.

An example of the code required, if a weight on Email were to be added is as follows:

SELECT WEIGHTS

GO TOP

REPLACE EMAIL WITH 60

SKIP 1

REPLACE EMAIL_NAME WITH 40

SKIP 1

REPLACE EMAIL_NAME WITH 25

SKIP 1

REPLACE EMAIL_NAME WITH 15

SKIP 1

REPLACE EMAIL_NAME WITH 25

SKIP 1

REPLACE EMAIL_NAME WITH 0

SKIP 1

REPLACE EMAIL_NAME WITH 1

SKIP 1

This code should be inserted between the following existing lines of code:

#ENDIF

DO UseTable WITH "PARAMS", "P", "", 1

Adjusting the Minimum Score to Report

To adjust the minimum score to report, from within the Job Editor screen double-click on the 'Options' cell of row 6 (where the 'Match Keys' cell is set to 'Change'), then when prompted; set the 'Setup Option' drop-down to 'Change' and click on 'Continue'. The main options for this row will now be displayed.

From within the 'Matching' tab of the options, set the 'Minimum Score to Report' setting to the desired score. Please note; this is set to 50 by default, to ensure that the Address level matches are found as well as the business and individual level matches.

Merge/Purge

Introduction to Merge/Purge

This section is only applicable if you have matchIT Plus or above.

The most typical use of Merge/Purge processing is when you have a file of existing customers and also an external list of prospective customers. Typically, you want to write to all the new prospects, but want to be sure you do not inadvertently send the same letter to any of your existing customers. In this case, you need to locate any of your existing customers that happen to be on the new list, and "purge" them from it.

If, on the other hand, you do want to send the same letter to all existing customers and prospects as well, what you can do is to create a new file (or table) by "merging" the new list and your file of existing customers, eliminating (or not creating), any duplication. These two operations have much in common, and in fact when merging, matchIT first purges records from the new file and then merges the remainder into the Main File. Thus purge is the same as merge, but stopping short of the final stage; no merging takes place.

Each input file must have been Imported into a matchIT Main File (using the Setup Wizard or Import records options) before you can do Find Overlap successfully.

You do not have to Find and Delete Matches from each individual file before running Find Overlap, but you must do so if you want the internal duplicates in each file removed!

The Basic Merge/Purge Process is as follows:

- Find Overlap.

- View/Verify the Overlap to check the matches found.

- Merge or Purge (for a suitable threshold score).

File Layouts when Merging Files

Before merging two files, the two Main File Layouts should be checked to pre-empt problems later on e.g. has one file got more address lines or wider fields? If so, make sure that the Main File is the one with more lines or the wider fields. If necessary, you should use Main File Layout from the Jobs/Setup menu to add or widen fields in the Main File.

If you are using the Multiple File Wizard to merge files, the wizard creates a master layout for you which preserves all of the data from the individual files i.e. it is a superset of all the individual structures.

The Main File is the first file you select. Purge Records does not change the Main File, but removes records from the Second File. Merge Files adds unique records from the Second File to the Main File.

Unique References

If you let matchIT allocate Unique References when the files were originally imported, you must make sure that the Unique References don't overlap between the files. To ensure this, you can set the Next Reference Number (in Matching Options) before Importing the second file, to a value at least greater than the number of records in the first file. Alternatively, you can set it to zero if you want matchIT to prompt you for the starting Unique Reference whenever you Import a file. If your Unique References are not unique between the two files, you can use Generate Unique References in the Fields menu in Tools/Database Utilities to regenerate them in either file or in the merged file.

If you are using the Multiple File Wizard, the wizard does this for you.

Find Overlap

Find Overlap matches the records from one Main File against records with the same key from another file. You can then view a report of the results and choose which records to merge or purge, individually or by a range of the scores. If you have several files to merge, you must merge each one at a time into the Main File, use Job Scripts or the Multiple File Wizard. If you want to remove internal duplicates in the files being merged, you must find and delete matches from each file individually, before running Find Overlap on the two files.

If you are using the Multiple File Wizard, the wizard does all this for you.

Obviously, to Find Overlap, two files must be open. If this is not the case, matchIT will prompt you for file names when "Find Overlap" is selected (from the toolbar or Merge/Purge menu).

Find Overlap

The "Matching Key and Range" screen is shown first, as for Find Matches. Choosing the keys for finding the overlap between two files is exactly the same as for finding matches within one file, except that the Match Keys displayed will be limited to fields which are in both files and have the same width. If for example, Zip is not displayed as a match key, it may well be because it has a different width in the second file from the first. To overcome this problem, you can type in a match key in the Index box of e.g. LEFT(ZIP,5). See "Match Key – Definition" to review the description of entering your own keys.

After matchIT has done its work and found any matches, a summary is displayed (in the usual report preview screen) and then the screen below, which offers the most common options:
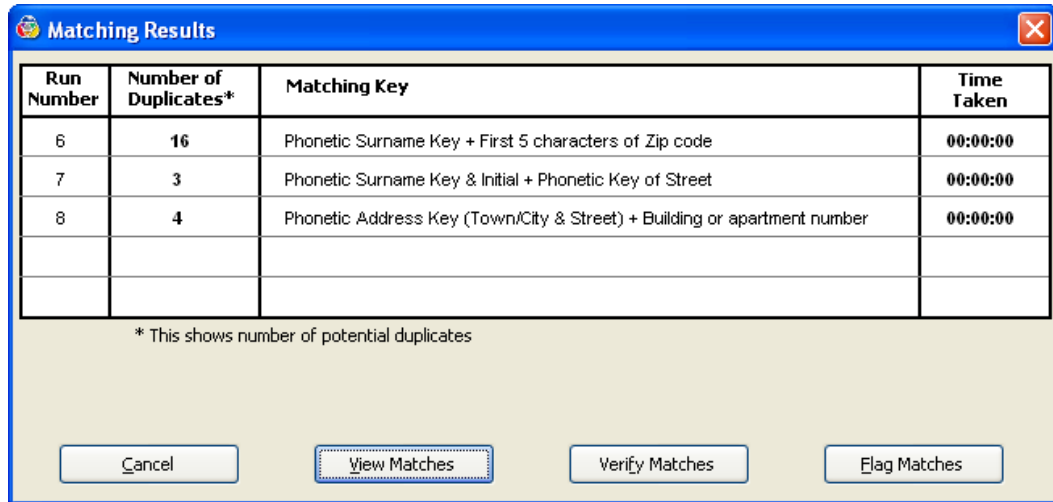
The choices from here are:

- View Overlap between the two files (similar to View Matches)

- Verify Overlap between the two files (similar to Verify Matches)

- Merge Files into the Main File

- Flag Records from the Second File

- Cancel

If you are selecting any of these options having previously closed your two files, ensure they are reopened in the same order as before. If you open them the wrong way round, matchIT prompts you to swap them, to ensure that it shows or processes the right pairs of matches.

As for Flag Records, when you select any option for processing the results of Find Overlap from the Merge/Purge menu, it allows you to view the individual match scores or score bands, before proceeding to the operation that you selected.

Overlap Summary

This is essentially the same report as for matching: see "Matching Summary" on page 85.

View/Verify Overlap

Viewing and verifying the overlap between two files is similar to viewing and verifying matches within a single file (see <u>View/Verify Matches</u> on page 86). The only difference is that different reports are used and you can only View Overlap in pairs, not sets.

View Overlap

View Overlap can be selected from the dialog presented after finding the overlap between two files or from the **Merge/Purge** menu.

Verify Overlap

As with **Verify Matches**, this can be selected straight after finding the overlap or from the **Merge/Purge** menu.

The screen displayed is the same as that for **Verify Matches**. For details see "<u>Verify Matches</u>" on page 87. If the two files have different layouts, the fields displayed for each record may be different. You can use the **Change Fields** button to change the order in which the fields are displayed, to help you compare the two records more easily.

Merge Files

When selecting Merge Files, you can enter the highest match score which you want to merge records for. If you enter zero, only those records on the Second File which were not given a Matching Score will be merged into the Main File. If you enter 100, only those records on the Second File which were given a Matching Score less than 100 in the MERGES table will be merged. If you enter all nines, or if the MERGES table is empty, all records on the Second File will be merged.

Flag Records

If you select Flag Records, you can enter the lowest score which you want to purge records for. If you enter zero, which gives the highest possible number of deletions from the Second File, all those records on the Second File which were given a Matching Score in the MERGES table will be purged from the Second File. If you enter 100, all records on the Second File which were given a Matching Score greater than or equal to 100 in the MERGES table will be purged. If you enter all nines, or if the MERGES table is empty, no records on the Second File will be purged. Finally, the Second File is selected as the Main File in use, to enable you to go straight to the Output menu without having to remember to select the right file first.

Transfer Data

This option in the Merge/Purge menu allows you to transfer data (such as telephone numbers, company data and geodemographic codes) from a master file to a customer file. You must run Find Overlap first.

The steps for this process are:

- Open your Main File and your Second File and run "Find Overlap" (from the Merge/Purge menu).

- Select "Transfer Data" (also from the Merge/Purge menu).

- Modify matching scores or click "Continue" (click Continue if unsure).

- The following screen will appear:

Data can be transferred between any fields which the Files have in common and these are shown in the left-hand panel. If you wish to transfer data from a field not listed, you must first use "Main File Layout" (on page 30) to add fields as necessary.

Either double-click or single-click and select "Move" or "Remove" to move or remove fields from the available list to selected field list or back.

Below the field picker there are other options can be set:

- **Transfer for scores** – data is only transferred for records whose matching score is equal to or greater than the value shown.

- **Transfer data** – from Main to Second shows where the data will be copied from to. Clicking on one of the buttons (main/second) will swap the order.

- **Always Replace/Only if Empty** – If there is data in a field that is about to be overwritten, matchIT needs to know whether to replace it (always replace) or leave it (replace only if empty).

- **Blank All Fields/Append to Existing Fields** – You can optionally choose to blank all the fields you are transferring to before starting, or append the data being transferred to that already in the field.

- In all cases, you should ensure that the field being transferred to is large enough to contain all the data or some of it will be lost. To make sure that only the best match is used to transfer data across where many records match one, define a field MATCHSCORE (numeric 8) in your receiving file, using "Main File Layout" (on page 30).

- Click "Begin Transfer" to start the process.

- matchIT will select this new table as the current one and a summary screen will be displayed.

Write Overlapping Records

This option on the Merge/Purge menu allows you to write overlapping records (after Find Overlap) from the **Main File**, to a new file e.g. to mail multiple buyers.

The steps for this process are:

- Open your Main File and your Second File and run "**Find Overlap**" (from the Merge/Purge menu).

- Select "**Write Overlapping Records**" (also from the Merge/Purge menu).

- Modify matching scores or click "**Continue**" (click Continue if unsure).

- A further window will appear, from which you select the minimum score for writing out the records. The default (as for Find Overlap) is to output records with a matching score greater than or equal to the Minimum score to report. Click "**Continue**" to continue.
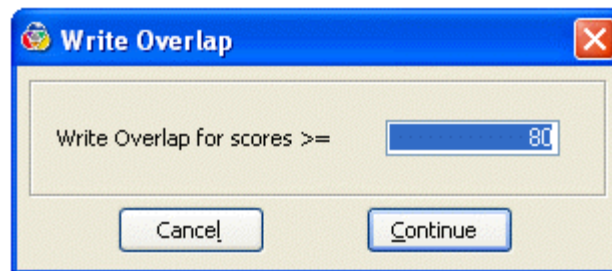
- You will then see window that allows you to either output the "**Overlapping Records**" or to access the "**Quality Assurance Wizard**". Select the "**Overlapping Records**" option and you will see a new window that will allow you to specify your new files layout.

- Select **Generate Output**, once your options have been specified, to begin the operation. You will now be prompted for a file name and location.

- matchIT will select this new File as the current Main File and a summary screen will be displayed so you can go on to produce output or return to the main menu (the **Done** button).

Job Scripts

Job Scripts Location

This section is only applicable if you have matchIT **Pro**.

To access the job script options, you can go to the *Jobs/Setup* menu, and then choose the menu option matchIT *Jobs*, followed by *Run/Amend Job Script*:
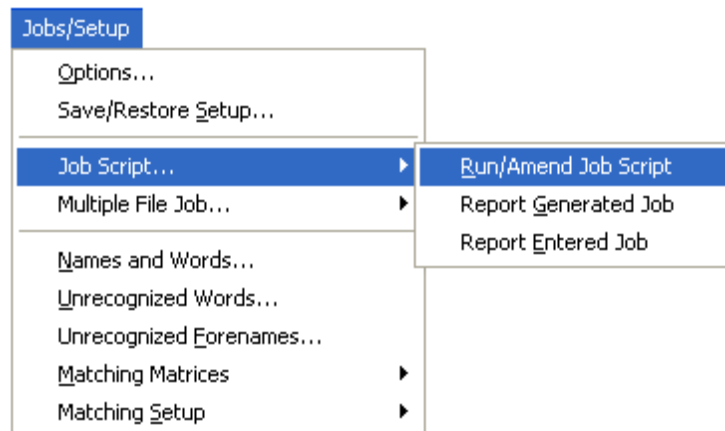
Alternatively, you can choose the Automation Wizard icon from the Wizard Pane.

Run a Job Script

You can also run any job by specifying a job name as a parameter of the matchIT program icon.

Job Scripts Overview

The Run/Amend Job Script option allows you to specify, run, clone and amend standard jobs consisting of multiple Import, Find Matches, Delete Matches, Find Overlap, and Merge/Purge steps. This option also allows you to specify an unlimited number of files to merge, with a pecking order, and any number of suppression files to compare the merged file with. You can also run FoxPro program files (either compiled FXP files or simple series of interpreted commands in PRG files) as part of the script, for example to copy a file, replace data fields etc.

Job scripts are very flexible and powerful automation tools for matchIT processing. Job scripts don't let you cut corners – they just allow you to repeat things easily. You still have to do the same amount of work specifying file names, Match Keys, deletion scores etc. – but with job scripts, once that is done, you can use that script again any number of times to repeat processing without having to specify it again. If you are not already fully familiar with the workings of matchIT, we strongly suggest that you become familiar with the normal menu-driven Import and Matching functions before you try and set up a job script. Once you have done this, the Multiple File Wizard is the easiest way to set up a multiple file job script and is also easier to use than the Merge/Purge menu options.

Job scripts may look complicated, but if you read and understand the information given here, setting them up and using them should be quite straightforward.

Job Maintenance

Selecting the Run/Amend Job Script option will bring up the Job Maintenance Screen, from which you will be able to create, delete or run job scripts:

The main fields on this screen are *Job Reference*, *Job Description*, and *Main File*, which are explained below:

Job Reference

The job reference is your identifier by which you define the job. matchIT also uses the job reference as a unique identifier for the job, therefore you cannot have two job references the same.

Job Description

The job description is an optional descriptive field to enable you to more clearly describe the job in question.

Main File

This field specifies the Main File for the job. The Main File is the first DBF file in your list of files, and it is the one which the other files are merged into (or the one Suppress files purge from). This is the file that will contain the results of any merge/purge processing when the job is complete.

You can choose any job by choosing from one of these three popups. If you choose a job, the other fields will be automatically updated to reflect your change.

Running and Resuming a Job Script

Below the main panel listing the available jobs are three buttons: **Start Job**, **Resume Job** and **Cancel**. Select **Cancel** to return to the main menu without running a job. Any changes you have made to the job scripts will be remembered. Select **Run** to actually start the selected job script.

Select **Resume Job** if you have had to interrupt a job script whilst it was running, and now want to resume from where you left off. Choose the step from which you wish to resume – it will default to the step which was last run, which is obtained from the Performance log (which you can view via the View menu, Performance Data, All Runs). Note that "step" does not mean line number of the job script; a single line may include, for example, one Import step, three Find Matches, two Find Overlaps, one Merge Files and one Purge Records.

Now we have explained how to actually choose and run a job script, we'll explain how to create them in the first place.

Creating a Job Script

To the right of the main panel are the buttons associated with editing the job scripts. To create a NEW job script, click on the *New Job* button.

You will now see a small window asking you for a new reference. Type in something here which will be the name of the job script. We suggest that you call it something descriptive. The only limitation on this name is that it must be a name you have not used before (i.e. it must be unique).

You will now see the Job Editor Screen, which is described below.

Modifying a Job Script

At the Job Maintenance screen, you can choose to edit an existing job script (instead of creating a new one which is described above). To do this, first choose the job you want to modify from the list of jobs available, then click on the Modify button on the right hand panel.

Whether you have chosen to modify a job, or create a new one, the next screen you will see is the **Job Editor**:
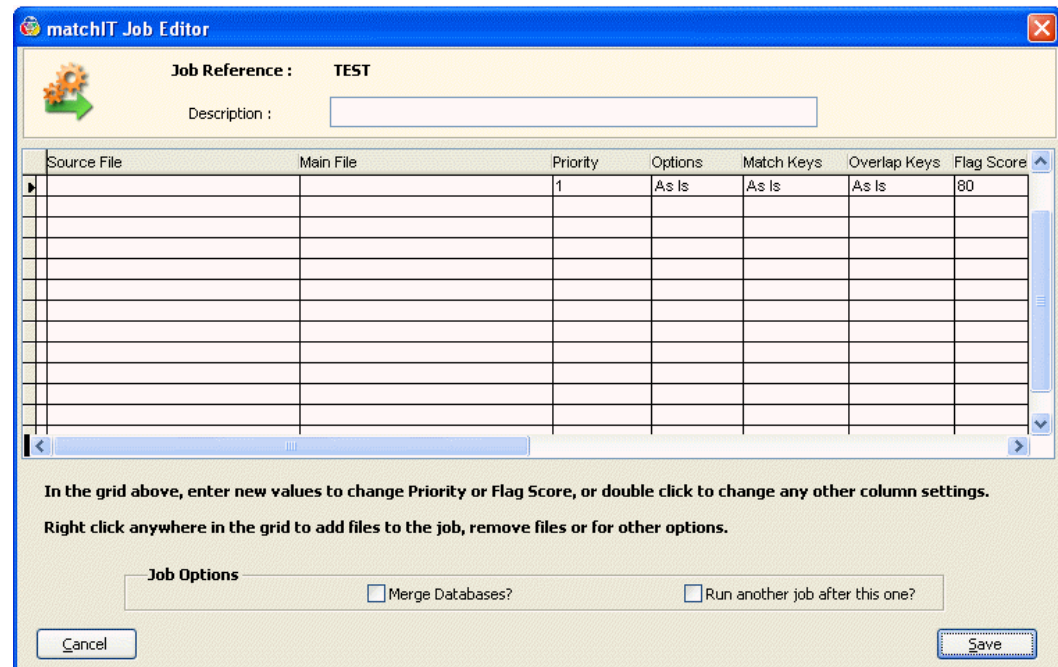
Explanation of Job Editor Layout

The main concept behind the job editor is that it is designed to represent a list of files to be processed. Each line will correspond to the processing done on one file. The processing is generally done in left-to-right, top-to-bottom order (like reading a book). Processing may include importing a source file (into a predefined Main File layout), followed by internally deduping it, followed by finding its overlap with the master file (the one at the top of the list), and the purging of duplicate records from the file, or merging the records into the master file.

Therefore, the columns correspond to these processes – first you have the **Source Name**, which will be the file name of the source text file, followed by the **Main DBF**, which will be the Main File layout that the data is imported into. This is followed by the **Priority** column, which indicated the files relevant priority in the list (low numbers = high priority). Next, we have the column for import & matching configuration, **Options**, and then columns for internal deduplication (**Match Keys**) and then finding overlap between two files (**Overlap Keys**). Finally, we have a column for the **Flag Score** to be used for processing.

In terms of file order, the first file is the **Master file**. This is the file that each find overlap is done against. Also, if files are merged, they will be merged into this file.

The files that follow this will be **Merge/Purge** files. Each of these files can be imported and internally deduped according to how you set up that line of the job, and then will be compared

with the master file. (Note: they will NOT be compared to each other). These files will then be purged of the overlapping records, or merged into the master file.

Following this comes **Suppression** files. These files can be imported, deduped and compared to the master in the same way as the Merge/Purge files. However, these files are special in that they contain **exclusion records**. That is, when the overlap is complete, the records will be removed from the master file, not the secondary file (which is the case if you select Purge Records from the Merge/Purge menu). Files such as the MPS (Mailing Preference Service) file are suppression files.

You can also add bespoke program files to the job scripts at any point. These files are not imported or processed in any way, but contain FoxPro program code. They can be used to insert special programming steps and options into the job scripts which would not normally be possible. A good understanding of the FoxPro programming language is normally required to create and use these kinds of files.

Therefore, a common job script could look like this:
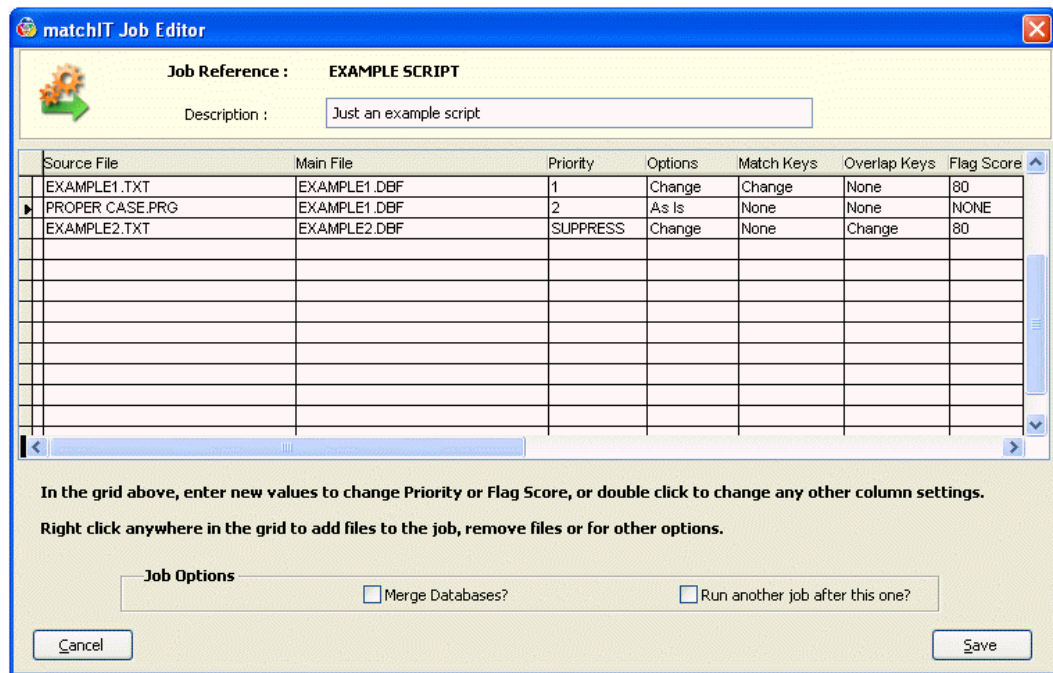
The following sections will explain the columns and settings above in more detail.

Adding and Removing Files (Rows)

To add or remove a merge file or a suppression file, you must *right-click* the mouse in the file listing part of the window, to see the menu. If you are removing a row that already exists, you must select that row by *left-clicking* on it first. Note that when you start editing a new job script, matchIT will automatically add a row for you.

To add a merge/purge file, click on the **Add Row** option. This will add a new blank row at the bottom of the list. Similarly, you can add a **Add Suppression Row** with the equivalent option. Note that when suppression files are added, they are always added to the bottom of the whole list, whilst merge/purge files will be added to the end of the merge/purge list only (i.e. before suppression files).

If you want to remove the currently selected row, click on **Delete Row**.

With this menu, you can also **Rename Job** (click on it and it will ask you for the new name – the same naming conventions apply as when you create a new job script). Furthermore, you can also **Clone Job** (copy the current job). When you do this, matchIT will again ask you for the new job script name. Note that when the copying of the job has been done, the job you will be editing will be the new job.

Explanation of Columns

Source Name

This column specifies the file name of the original data – the data before you have done anything to it. This is the file which will be imported into matchIT.

If you double-click on an entry in this column, you will see the following screen:
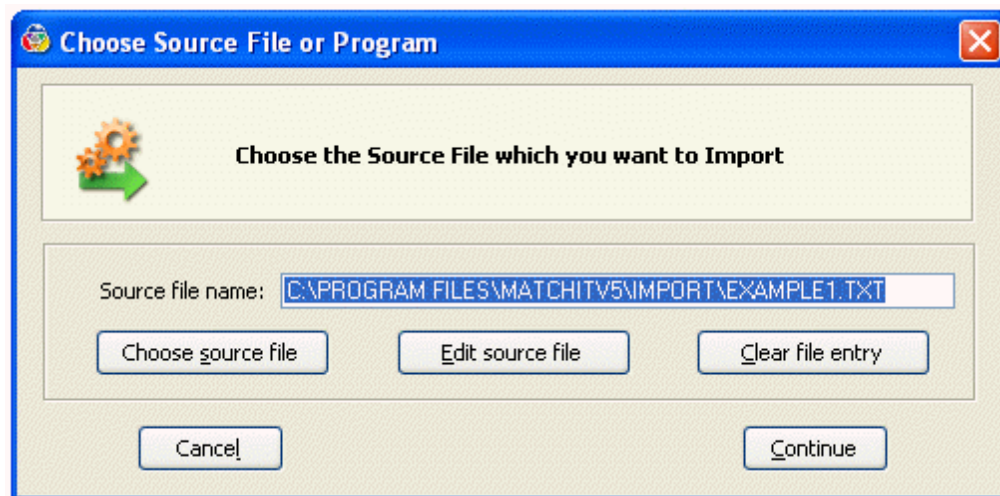
You can either type directly into the field to change the file name, or you can click on the Choose source file button to choose another one. Click on Edit source file to view (and perhaps change) the data file.

Special Cases

There are two special uses for this screen that the user should be aware of:

- If you choose a file with extension .FXP (a Visual FoxPro compiled executable file) or .PRG (a text file which contains a FoxPro program – see later for more details), then this file will not be treated as an import file, but as an **executable** program file. This means that matchIT will not process the data in the file, but will attempt to run it instead. In this case, all the other columns in the row are irrelevant, and will be ignored by matchIT. Note that compiled Visual FoxPro programs must be compiled using the same version of Visual FoxPro as your version of matchIT – you can ascertain this via the Help menu, About matchIT.

- If you have an import format other than EXT or Regenerate Keys (see Setup below), and the source file name is blank, then matchIT will not do any import processing on the file. matchIT will assume the file has already been imported, and will not try to do so again. In the case of the EXT format, the source file should **always** be empty. (This format means do the **generate keys** step only – don't add records from a source file previous to this). For more information on this format, please consult the Import Options section.

Main File

This column corresponds to the Main File that will contain the data after it has been imported. To choose a file for this, just double-click on the field, and matchIT will ask you for a file name. Note that the file must exist before the job script is run – the job script cannot automatically create Main File layouts. The File can of course be empty. When the job is run, any data in this file will be over-written by the data in the source file.

Priority

The priority of a file corresponds to its place in the processing order. The lower this number, the more important a file is. The file with the highest priority (usually 1), will be at the top of the list. This file will be the Master File, with which all other files are compared, and into which files can be merged.

Priorities are automatically assigned when you add merge or suppression files, in the order that they were added to the job. You can change any of these priorities by typing into an appropriate number into this field. This field supports decimal numbers if necessary.

For suppression files, these will have the priority set to "SUPPRESS". You can make a merge file a suppression file by typing in the word SUPPRESS into this field. Similarly, you can turn a suppression file into a merge file by typing a number into this field.

Suppression files do not have priorities amongst themselves, as they only act to remove records from the master file (when a record is removed from the master file, it doesn't matter if suppression file A or suppression file B removed it – the result is the same).

If you blank out the priority field, this will be treated as priority zero by matchIT. If you have two files with the same priority, the order they will be processed in will be the order that they were added (i.e. the one you added to the job first will be processed first).

Options

The *Options* column is where most of the matchIT processing options are set. These include options such as operation parameters and Import Options – please consult the matchIT Options section for more information on the options available.

Double-clicking on a row in the Options column will give you a choice of three options:
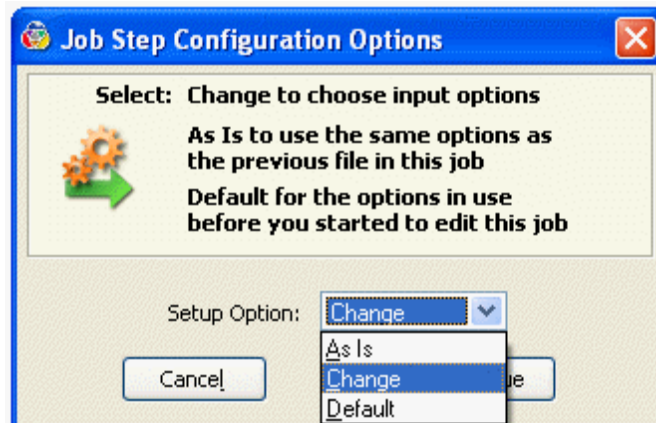
Change

Select *Change* to view the standard matchIT Options screen (with a couple of tabs that do not relate to job script processing removed). From here, you can change processing options as required.

As Is

This option has two slightly different meanings, depending on whether it is chosen for the Master File or any other file.

If *As Is* is chosen for the Master File, this means that the matchIT options chosen will be the same as the current configuration. (Check this via the Jobs/Setup menu, Options before starting the job script).

If *As Is* is chosen for any file other than the first one on the list, the job script processing will take the matchIT configuration from whatever was set up in the previous line in the script. (That is, it remains unchanged).

Default

Whenever this option is chosen, the job script processing will take the configuration from the current setup of matchIT. (Note that this has the same effect as *As Is* for the first file in the list).

Notes

One pitfall for the unwary in job script configurations is the *generation of unique reference numbers*. As has been stated elsewhere, it is vital that within a single file, all unique reference numbers *remain unique*. This is simple within only one file, but when you are cross matching and merging two or more files together, then maintenance of the reference number becomes more complicated.

One way to resolve this problem is to assign a new unique reference starting number for each file that is being imported. You can do this by setting the *Options* column to *Change* for every file that is being imported, and by manually setting the starting reference number in the matchIT options for each file. Alternatively, if you set the job script up using the Multiple File Wizard, it will ensure that the unique references are really unique across all the records in all the files in the job – otherwise it will regenerate them starting from 1.

Another feature of the job script processing that can be used is the *Input Format* option. As mentioned above, if you set this to anything except *EXT*, and have a blank *Source File* name, then no import processing will be done. This can be very useful for quicker processing of files that have already been imported. Furthermore, if you do not want to add new records into a file, but just want to regenerate the matchIT fields, you can set the Input Format to EXT.

Match Keys

Match Keys are used for the internal deduplication step of matchIT. As with normal use of matchIT, you should specify a set of Match Keys in order to carry out the deduplication. This corresponds to the Find Matches menu option.

In a similar way to the Setup column, double-clicking on this field will show you an options screen:
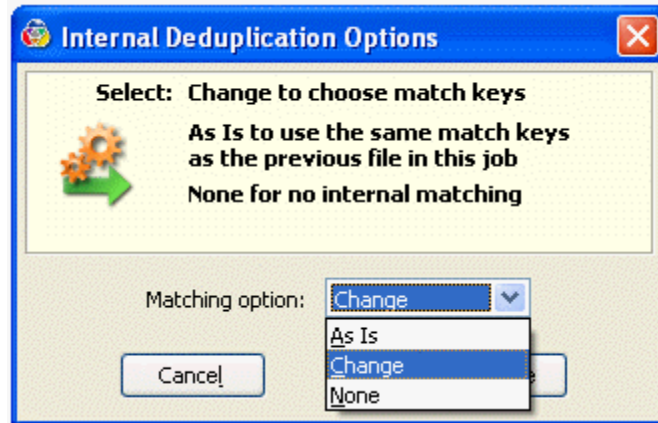
Change

This option works in an equivalent way to the Change option in the Setup column. Choose this option and click on Continue to go to the Match Keys screen where you can edit the Match Keys that will be used in the find matches step in the job.

As Is

Again, this works in a similar way to the As Is from the Setup column. Selecting it means that the Match Keys used for this file will be the same as for the previous file in the list.

Note that this option does not apply to the first file in the list.

None

If you select this, then no internal Find Matches steps will take place. Use this for files that have already been deduped by matchIT, or files which do not need to be cleaned.

Setting the options in this column will only cause matches to be **searched for**. If you want to actually delete the matches as well, you must set the Deletion Score to a non-zero value.

Overlap Keys

This sets the keys that will be used to find overlap between the master file, and the current file. This works almost exactly the same way as the Match Keys option works.

The only real difference is that the first file can only have this set to None, as it can't find overlap with itself.

As for the Match Keys, this will only **search for** the overlap, it will not actually process the results – so if you want to merge the file with the master file or purge records from this file, set the deletion score column to the appropriate score value.

Deletion Score

This determines the score at which a potential match is considered to be a true match. It is used in two ways. Firstly, for internal deduplication, this is the threshold score for deleting records. Secondly, when a Find Overlap is carried out, this is the threshold for purging records (usually in preparation for merging them into the master file).

With the master file, this score only refers to the deletion threshold for internal duplicates.

If you do not wish records to be deleted, set this value to zero (0) or "NONE".

Other Job Script Options

Merge Databases

If you check this option, each file after the first one will be merged with the first file. Records that have been found to overlap with the master file (via the Find Overlap steps) will not be added to the master file.

If you set this option off, then the overlapping records in the "merge files" will be purged according to the deletion score in each file, but the file will not be added to the master file. Note that if the deletion score is set to zero or NONE, then records will not be purged either.

Note that the Multiple File Wizard is an exception for use of this option, as it always merges the files into the master file irrespective of whether this option is set on.

Run Another Job At End?

Sometimes, you may wish to link more than one job together. This is especially the case when running small program files, or when you want to merge files into multiple master files. In all cases, the first job will be totally complete before any part of the next job is started.

To enable this option, click on the check box. You will be presented with a job script picker screen like this:
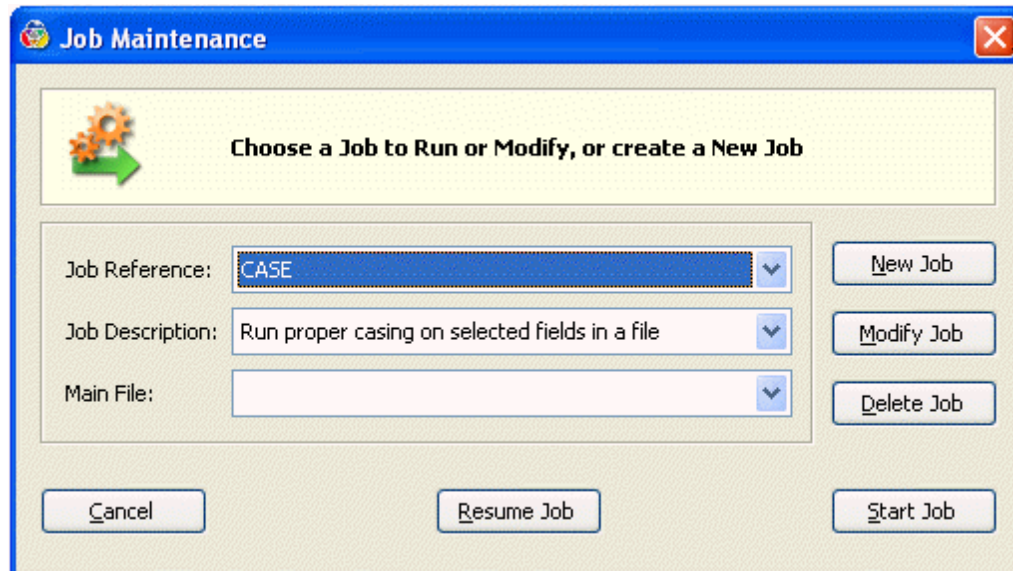
To run another job at the end of the current job, just choose the new job from the list available. Once you have chosen the job you want, click on the Continue button to confirm the choice.

If you don't want to run a job at the end, click on the Clear button. This will tell the job script processing NOT to run a job at the end of the current one.

Click Cancel if you are happy with the current state of this option, and don't wish to change it.

Running a Job Script automatically when you start *matchIT*

You can run a job script automatically when matchIT starts by modifying the properties of the matchIT icon so that after the *Target* you add /R: followed by the job script name e.g.

"C:\Program Files\matchITv5\matchit.exe" /R:jobname

where **jobname** is the reference of the job script. To quit matchIT automatically at the end of the job script, you have to run a program at the end of the job that simply contains these commands:

ON SHUTDOWN

QUIT

Multiple File Wizard

Preparation for using the Multiple File Wizard

The **Multiple File Wizard** enables you to easily set up a multiple file merge or merge/purge job. Before using the Multiple File Wizard, it saves time if you set the File Locations for the Database and Import directories to those containing the files for the job (from the Jobs/Setup menu, Options.)

There is a limit of 2Gb (2,147 Mb) on the size of any Main File.

Please note that all the files in the job are merged into a single master file, which the Multiple File Wizard creates, containing all of the fields and all of the records from any file within the job. There is a limit of 2Gb (2,147 Mb) on the size of the resultant master file, which in practice could be about 5 million records – the actual number of records depends on the record length i.e. the number and width of fields in the resultant master file layout. If you want to process more records than this, it will usually be because there are some very large suppression files involved. In this case, you should use the Multiple File Wizard for the mailing files and any small suppression files, then use Find Overlap between the large suppression files and the resultant master file. This also reduces the processing time significantly, because matchIT does not need to process match keys for all of the records within the large suppression files – obviously the Multiple File Wizard does not report internal matches within any suppression file, but it will take significantly extra time if very large suppression files are built into the master file. If you want to use Find Overlap for such a job but still include the suppression files in matrix reports from the job, please consult your support provider for job scripts that achieve this result.

Multiple File Wizard: Step 1

Select the **Multiple File Wizard** option from the matchIT Wizard Pane or from Jobs/Setup, Multiple File Job menu. You will be asked for a Job Reference, and a Job Name for the resultant reporting. After you have entered these details, select Continue.

Multiple File Wizard: Step 2

Then you will get the options to configure the first file:
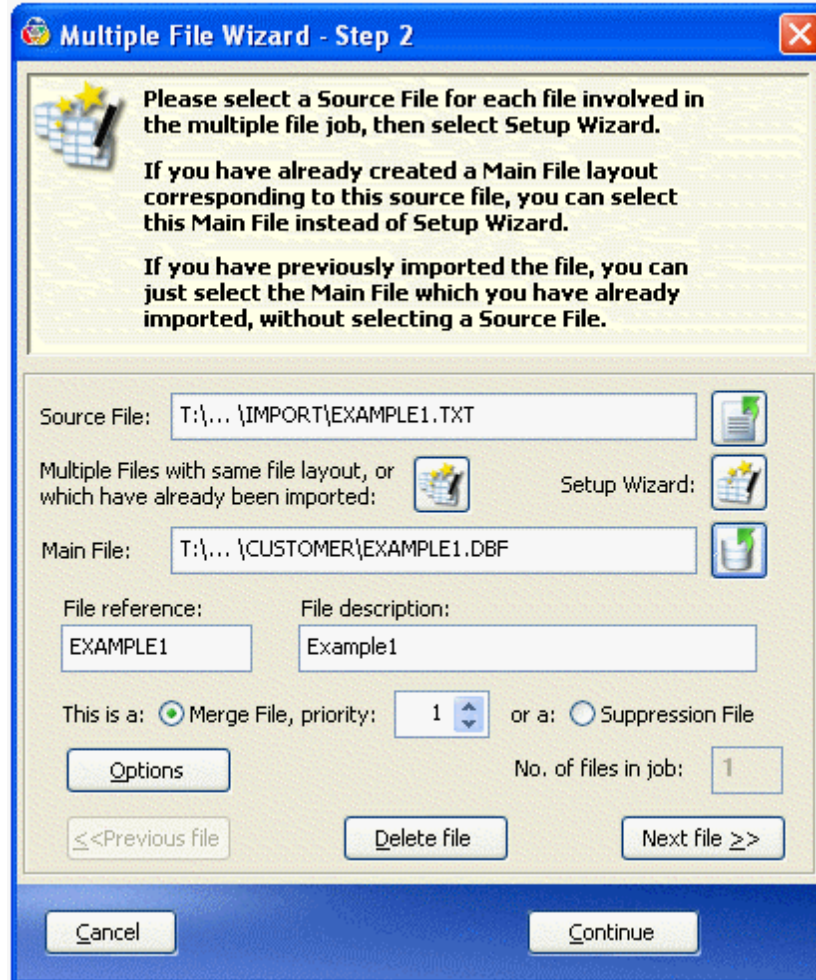
For each file, perform the following steps:

- Select the source file that you would like to use. If the file has already been imported into matchIT, there is no need to specify a source file.

- If the source file has been imported into matchIT specify that DBF as the Main File. If you have a "shell" DBF to import your source file into, specify that DBF as the Main File. If you do not have a DBF, you can use the Setup Wizard button to create a "shell" DBF for importing. When using the Setup Wizard option the DBF that is created will automatically be set as the Main File.

If you have several files with the same "exact" layout, you can utilize the "Multiple Files with same file layout.." button. From the next window, you can specify either several source files or several DBF's that have already been imported. You can also specify the priorities and Suppression files from this window.

- Select Options and set the Input Format to the format of the source file e.g. SDF for fixed width files with Carriage Return, COMMA for comma delimited.

Multiple File Wizard

If the file already has list codes and priorities in, these will be disregarded. If they are labeled LISTC and LISTP, they will be replaced by the list code and priority you enter on this screen.

- You can also change any of the other Input Options e.g. Proper Case Incoming Data, Addressing.

- If the file is a suppression file, select the Suppression File radial. If it is not a Suppression File, make sure the priorities are as you want: 1 is the highest priority i.e. the file for which you most want to keep records where they match records on another file. If you enter the highest priority file first and then in order of priority, the priority defaults to the right value.

After you have selected all the files, select Continue.

Multiple File Wizard: Step 3

You will then see the window below:
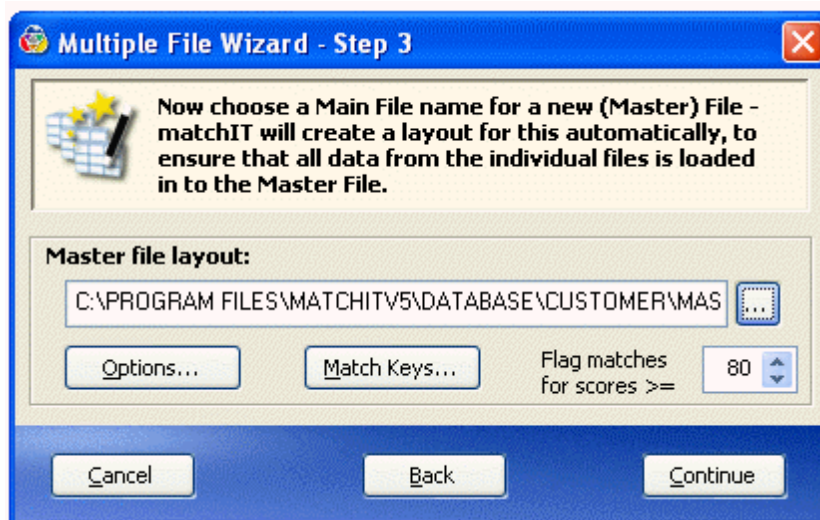
Click on the file selector button⬚ and enter a name and path for the Master file. matchIT will create a merged table layout with all the fields from all of the individual files, with the maximum field width from each file. All input files (including suppression files) are merged into the Master File. Each file is automatically given a list code reference (in the LISTC field) so splitting out the non-duplicate records from each original file is possible after the job. You are prompted to remove the Suppression records at the end of the job.

Then choose the Match Keys that you want to use, as for Find Matches. You must also enter a deletion score so that the job will automatically delete the matching records (as in Delete Matches).

127

Multiple File Wizard: Step 4

Finally, you will see the screen below:
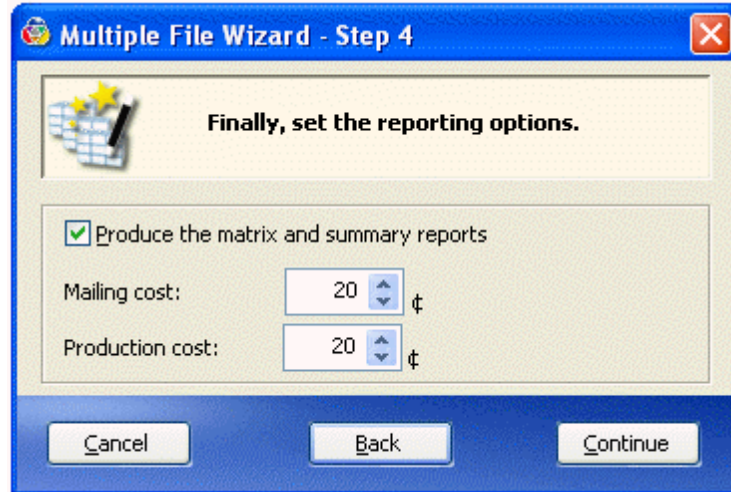
If you want the matrix and summary reports produced, check the option shown and enter the costs.

Run/Amend Job Script

You will then be prompted as to whether you want to run the job immediately. Whether you do or not, you can amend and run the generated job at any stage from the Run/Amend Job Script option on the matchIT Jobs menu or from the Automation button in matchIT's Wizard Pane – your job is saved under the Job Reference name you input in step 1. This is described in the Job Maintenance help topic.

You can change the generated job script later to change file names and file setup options, but you must not change processing priorities or add files, because there is currently no way to specify the list code and list priority for the matrix report in the Run/Amend Job Script window.

You can print the Data and Matching summaries and the matching reports from the Input and Matching menus as normal, with the Master File selected. If you change your mind about the Deletion Score you used, you can undelete records using Database Utilities (matchIT's program folder), Verify Matches to determine the new deletion threshold, Delete Matches and finally Produce Matrix reports from the Jobs/Setup, matchIT Jobs menu.

If when using Verify Matches you want to interactively decide on true and false matches in a "grey area", you can flag any pair as a False Match, or increase the score of any pair so that it is included in automatic deletion. You must NOT delete matches interactively in Verify Matches if you want the Matrix Reports to tally.

If you do not remove the suppression records from the Master file when prompted, or you undelete them, you can remove them later using the Delete Suppression Records option from the Jobs/Setup, Multiple File Job menu.

Produce Matrix Reports

The 'Produce Matrix Reports' enables you to output the matrix and summary report for a multiple file dedupe job at a different time from running the job script, or generates these reports from a single file dedupe job that has been configured to be compatible with a multiple file job. Note, if you generate reports from a file rather than a job script, then the long descriptions of the constituent files in the job will not appear on the summary report. Instead the report will say 'Mailing File 1', 'Mailing File 2', and 'Suppression File 1', 'Suppression File 2', etc. This is the only difference in the reporting.

Choose 'Produce Matrix Reports' from the Multiple File Job sub-menu of the Jobs/Setup menu. You will then see a screen, which gives you the choice of choosing which job script you want to produce the matrix reports for, or which single file. If no multiple file job scripts have been set up (that is, using the Multiple File Wizard), then the first option will be disabled, and you will only be able to choose a file. Once the job script or file has been chosen, and the OK button selected, you will then see another screen, which gives you reporting options. You can choose to regenerate the reports here, which you will need to do if you have generated reports for another job or file in the meantime, or need to change the mailing and production costs. However, if you have generated the reports for this job or file previously, and you are happy with the mailing and production costs, then you don't need to regenerate the reports. Clicking the 'Continue' button will then allow you to choose an output format for the matrix report and the summary report.
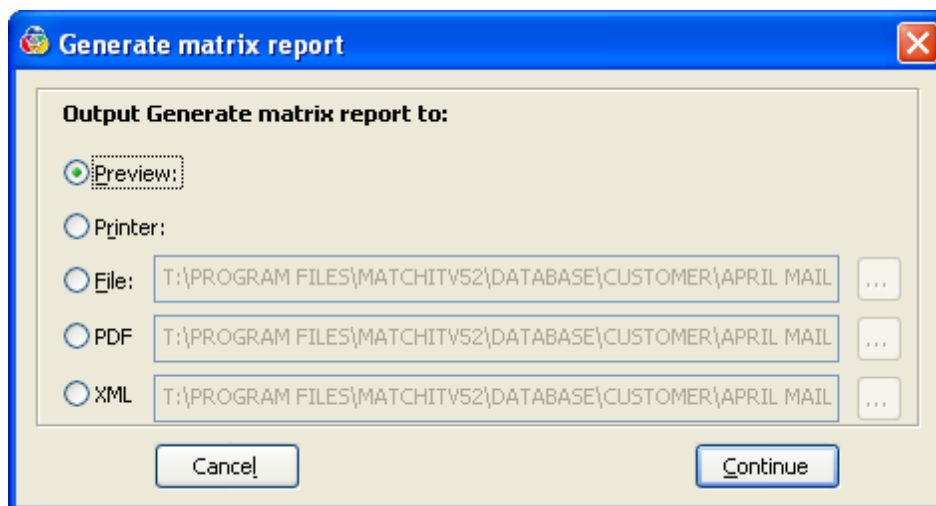
Output

Producing Output

You can choose from File, Label and Page options: File creates a file on disk to send elsewhere for printing, Label prints mailing labels and Letter/Page/Report prints letters or envelopes according to a layout that you establish via the Output Layouts option. The information that you enter is mostly similar for all three options.

Main Options

Layout name

This will either show "Use all fields" or a predefined output layout, which you can change by clicking on the ▣ button. You can edit the output layout selected – see Changing Output Layouts on page 133.

Defaults:

File Layout defaults to "Use all fields".

Label Layout defaults to ADDRESS.LBX, which is a standard layout which should be suitable for most mailing labels, in terms of content. To change the label dimensions and the number across or down the page, use the Label Wizard via the Edit, Output Layouts menu.

Page Layout defaults to MAILSHOT.FRX, which you will most likely need to amend.

Destination File Name

The name of the output file.

Output Format

This shows the current output format, which can be changed for output to file by using the drop-down list. The options here are basically the same as for Input Format except for:

- Paradox format (which is not available for output)

- Access (which you can only output by using an ODBC connection, although you can output as a DBF file which you can then read easily in Access, including field labels)

- Excel (which is limited to Excel 5 format and thus 65,000 records, unless you use an ODBC connection).

For ODBC format, the Connection File Name must be a pre-existing connection file (see "ODBC Manager" on page 39). You will be prompted for a file name.

Destination

If the output is labels or letters, it can be previewed on the screen or printed directly.

If you are outputting to file, this option cannot be changed.

Sequential Numbering

- **Sequential numbering**: check this option to generate sequence numbers in the output.

- **Record at which to start printing**: enter the number of the record from which you want to start the output.

- **And number it**: enter the sequence number you want to output, in the first record output (only available once sequential numbering is selected).

- **Number of records**: Either choose "all" (the default) or enter the number of records you want to output.

Shuffle Up Address Lines?

Check this option to close up the address lines to eliminate any blank lines in the output. matchIT will prompt you to say whether you want to close up the zip code field as well.

Use Campaign History Options?

Check this option to use Campaign History functions (only if you have matchIT **Campaign**).

Include Flagged Records

Include records that have been flagged, as duplicate or otherwise, in your output file.

Sampling Options

You can print a sample of the records by checking the box and entering a number. e.g. 1 in 1000 sampling will output every 1000th record in the file. See also "Quality Assurance" on page 138.

Filters and Ordering

You can create Filters and/or Indexes for any of your Output Files. For example, you may have a filter of **COUNTRY='US'** and an index of **SURNAME+FORENAMES**. This would output all records with a Country value of "US", the resultant file would then be sorted by Surname then Forename.

Campaign History Options

These options are only available in matchIT **Campaign**.

Campaign ID

This is stored in the MAILING_ID field for each record, if this field is defined in the Main File Layout. Mailshot letters in the delivered MAILSHOT.FRX template also take their standard text from the text record for this campaign.

Date

This is the date printed on Mailshot letters and the date stored in the MAIL_DATE field for each record, if this field is defined in the Main File Layout. This feature is available only in matchIT **Campaign**.

Use Reference File

If you have a file of unique reference numbers which you want to mail, one reference number per line, then you can specify the name of the file and matchIT will extract those records from the Main File and output just those records. This feature is available only in matchIT **Campaign**.

Changing Output Layouts

This section is not applicable if you have match11 **Lite**.

Selecting a Layout to Modify

The "Output Layouts" option from the Edit menu allows you to modify file, label or letter/page/report layouts. You can create new formats by saving modified ones under a new name, or by selecting New when asked for a layout to modify.

Select which type of layout you wish to modify and then select the layout name.

File

The list of fields for output to file can be modified and saved for future use. The top line in the dialog shows the file name and path of the current layout: click on the ☐ button to choose other output layout files (*.OPL). Choose the Edit button to change the output layout. You can also select this option from the Edit menu, Output Layouts.

You will then see the File Layout Designer:
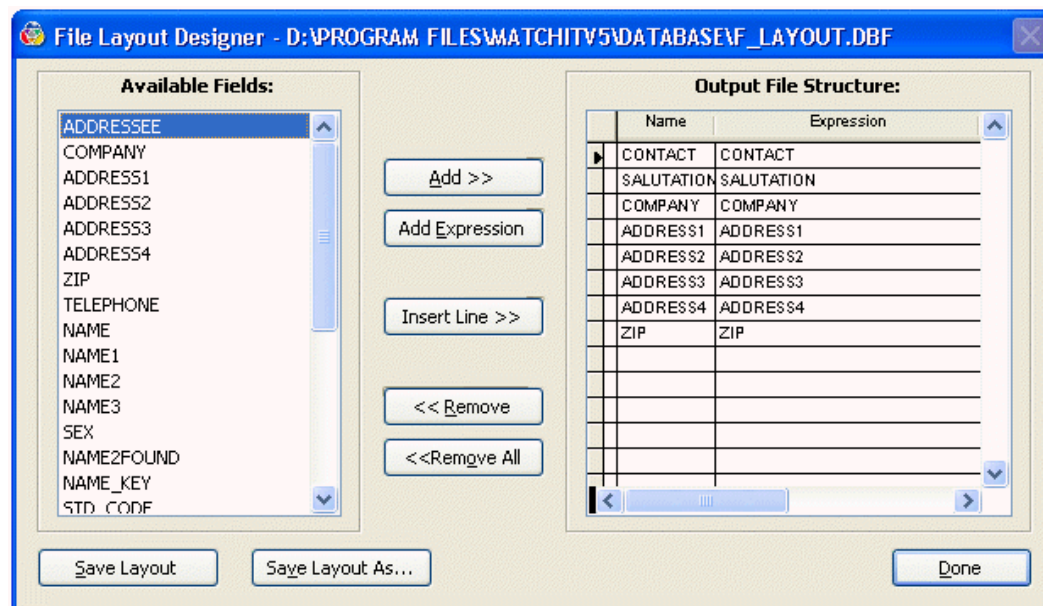
This allows you to specify which fields (and field expressions if you wish) will be output to your file. The field picker shows you all the fields which are available in the currently selected file in the left hand pane, and all the fields which have been selected for output in the right hand pane. You can click on the *Remove All* button in the middle to remove all the fields selected.

You must select the fields from the left hand pane in the order that you want them output. You can scroll down the available fields list to see more fields and double click any field name to move it to the right hand pane. You can also add or insert Field Expressions into the layout e.g. PADL(ALLTRIM(SEQ_NUM), 7, '0') to output a sequence number padded out with leading zeroes (you must output the normal sequence number via the check box as well, because this is the SEQ_NUM field).

Click *Save Layout As* when you have finished editing your layout(s).

Label

Mailing Labels are address labels, the layout of which can be modified via the Label option of this menu. New Label layouts can be created using the Label Wizard option.

The Label Wizard

- In a few steps you can create a simple label which can be used or further modified:

- Select Edit -> Output Layouts -> Label Wizard . If you do not have a Main File open, choose one now.

- matchIT asks a preliminary question about your printer, as it does not support Dot Matrix and line printers.

- Next, the Label Dimensions window is displayed:
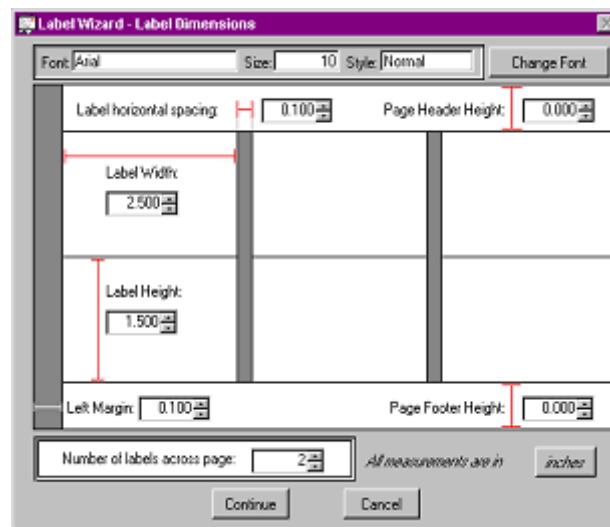
- Enter information about label dimensions, margins, and number of labels across the page. You can also select Change Font to change the font used and whether the dimensions are in inches or centimeters.

- Click "Continue" to move to the next stage.

- Now you can select the fields you would like on your label. Check the options required and change the spinner at the bottom to reflect how many address lines are to be used (up to a maximum of the number in your file).

- Click "Create Label" when finished and save your new label.

- You will then be shown your new label in the Label Designer window for further editing. This activates a new set of menus. To display the toolbars, you must select these options from the View menu. Context-sensitive help is displayed in the bottom left-hand corner of the screen.

- You can move the fields around by selecting the box containing the field, then drag it to another place.

- You can add new fields by firstly viewing the Report Controls Toolbar (View menu), then clicking on the abl button.

- When you now hover the mouse over the label it forms a crosshair. Click and drag a box to anywhere where there is space.

- An expression builder will pop up. Type the field name you wish to include on the label.

- You can preview the layout without leaving the design view, by selecting Preview from the View menu.

Defaults

Labels should show the direct or residue selection code of their destination, as well as the name and address. The default label layout for both these types of label is ADDRESS.LBX, which is a standard content that should be suitable for most mailing labels, including a person's name, job title, company, 5 address lines plus state and zip. Any fields which are not defined in your Main File are not printed.

Letter/Page/Report

matchIT's reports as well as your own letter layouts can be edited to include any fields or wording you need. They all use the same report designer as the Label Editor, described above.

Defaults

MAILSHOT.FRX in the Report sub-directory provides a model for printing "mail-merge" type letters. CLIENT.FRX in the same directory is a model for a checklist of customer names and addresses.

There are standard reports provided for printing residential and business records in pairs or sets, namely RESPAIRS, RESSETS, BIZPAIRS and BIZSETS. These formats show the person's name in the format in which it was imported, zip, 4 lines of address and (if present) the telephone number. The business formats also show the name of the business after the person's name. If you want to show more or less data, vary the order of the fields or the field widths, you will need to create new report format that you can base on the closest standard report by modifying the standard report and saving it under a new name.

In order to view or change the printing format of any of the reports, or to create new formats, or to change the type of printer, select the Output Layouts option of the Edit menu.

Text

With this option, you can create and modify freeform text for a letter easily, but this text can only be printed in a consistent typeface, not bold, italics etc for individual words or phrases. You can also enter a campaign ID and date, to be stored in fields labeled MAILING_ID and MAIL_DATE in the Customer records – this feature is available only in matchIT **Campaign**. If they are not defined in your Main File Layout, the information is not recorded.

Changing the Printer Setup

Use the Windows Print Manager to control printer selection and setup.

Changing Report Layouts

The Format and Report menus allow many different operations. These include:

- defining new fields and deleting existing fields

- changing field lengths and positions

- changing, moving, adding or deleting headings.

You can cut and paste field definitions from one report to another. You can also move or cut and paste groups of fields.

Don't change the field expressions DUP_REF(REF1) and DUP_REF(REF2) in the matching reports, or move them from their position at the beginning of each line (except that you may position them after the Match Score field). If you do, you may find that data from the wrong client records is printed in the report. You can correct this problem, if it arises, by using the Report menu option "Send To Front".

PRINT format output

Please see the Online help for more information.

Output Flagged Records

This is an option from the Output menu. It is used to create a File consisting of all the flagged records from your Main File.

To create a File containing the flagged records, follow these steps:

- Select "Output Flagged Records" from the Output menu. If you do not have your Main File open, you will be prompted to open it.

- A window will open displaying several different output options, including an option for output format. Once you have set your options e.g. supplied fields or all fields, select Generate Output. You will then be prompted to save the flagged records file.

matchIT will now copy the flagged records from your Main File into the new File you just named and unflag them from this new file. When it has finished, matchIT will automatically select the new file as the Main File.

Quality Assurance

Quality Assurance Introduction

We **strongly recommend** that you use appropriate Quality Assurance Checks to check all output from matchIT before you use it in a "live" or production environment e.g. before you print address labels. matchIT is a very flexible system which will handle all sorts of data, but this very flexibility makes it impossible for us to check out all combinations of parameters, data and user actions in our testing. The duty of care in checking output has to lie with you, the user.

To assist and complement your Quality Assurance Checks, matchIT provides several features.

These procedures will take quite some time on a large file, but the potential for large quantities of errors is much greater on large files, so it is advisable to take the time required to follow these processes.

Please let your support provider know of any errors, omissions or any unexplained results, so that we can address any problems that arise.

We will update this section of the manual as we get feedback and suggestions from our users about Quality Assurance methods, so for the latest version of this topic, please consult the Online Help.

Import Results

This option, from the Import menu, gives you options to preview and print the **Data Summary**, **View Records by Category** and **Non Alpha Numeric Characters**. The Data Summary is produced automatically at the end of Import (or Find Matches if you have Wait for User after Import switched off). You should examine this to make sure that the volumes of data and the nature of the data shown on the summary conform with what you expect e.g. with the last time that you ran similar data through matchIT. The Data Summary is regenerated each time you ask for it, so will show different values after deletion of matches than those generated immediately after Import.

View Records by Category allows you to "home in" on many of the categories of record for which entries are shown on the Data Summary and browse them on screen, print them, delete them from the file and/or copy them out to a separate file.

View Records by Category

You should also use **View Records by Category** from the **Import** menu or the **Quality Assurance Wizard**. It allows you to browse all records meeting selected criteria, most of them are categorized on the Data Summary that is displayed after Import. You can change records and delete records and those changes will be retained in the DBF file – so you can make any corrections that you wish in this view.

You can select several categories to view at once. You can also choose to delete the selected records, print them and/or to write them out to a separate file.

After you browse the selected records, matchIT returns you to the same screen to select more categories if you wish. To leave the screen, select **Close**.

View Data

"**View Data**" from the **View** menu allows you to browse records in different orders e.g.

- browsing in Prefix or Salutation order will show groups of records, typically at the top or bottom of the view, which have default salutations or unusual titles. The vast majority of records will be in a contiguous block i.e. MISS, MR, MRS, MS and Dear Miss, Dear Mr, Dear Mrs and Dear Ms – you should at least spot check these, but then look in detail at the records at the top and bottom of the list (i.e. not Miss, Mr etc.), because here in particular you may find errors or unusual salutations. You will also find default salutations that perhaps should not be default

- browsing in Address line 1 order may show undeliverable items at the top

- browsing in Company order will show blank companies at the top

Each time you select to browse in order of any field for which an index was not created at the end of Import (such as all the fields mentioned above), matchIT will create an index – this will take quite some time on a large file, but not as much time as sorting out an error when it's too late to pre-empt it.

You can change records and delete records in View Data and these changes are retained in the DBF file – so you can make any corrections that you wish in this view.

Sampling

On the Output Options screen you can select One in N sampling, or print/output a range of records.

One in N sampling

One in N sampling from the Output Options screen is available to help pre-empt problems with final ("live") output. When you choose this option, an intermediate table is created containing a selection of the records in your Main File. For example, if you have a Main File of 100,000 records, and choose One in 1000 sampling, your test file will contain 100 records – the 1000th, 2000th, etc. from your Main File.

Obviously, sampling does not guarantee that there will be no errors in your data.

Other Sampling

Alternatively, on the Output Options screen users can choose to output a number of records, the first 100 records from the File, or the 1000th through to the 1100th record.

Preview Output

You can also preview output using matchIT's default address labels, even if you are not printing labels from matchIT or printing labels at all, because this is an easy way to visually check name and address data before committing it to any further processing.

Quality Assurance Wizard

The Quality Assurance Wizard will guide you through some basic Quality Assurance steps. The Quality Assurance Wizard is also available from the dialogs displayed after Delete Matches, Merge and Purge.

Quality Assurance Wizard – Input
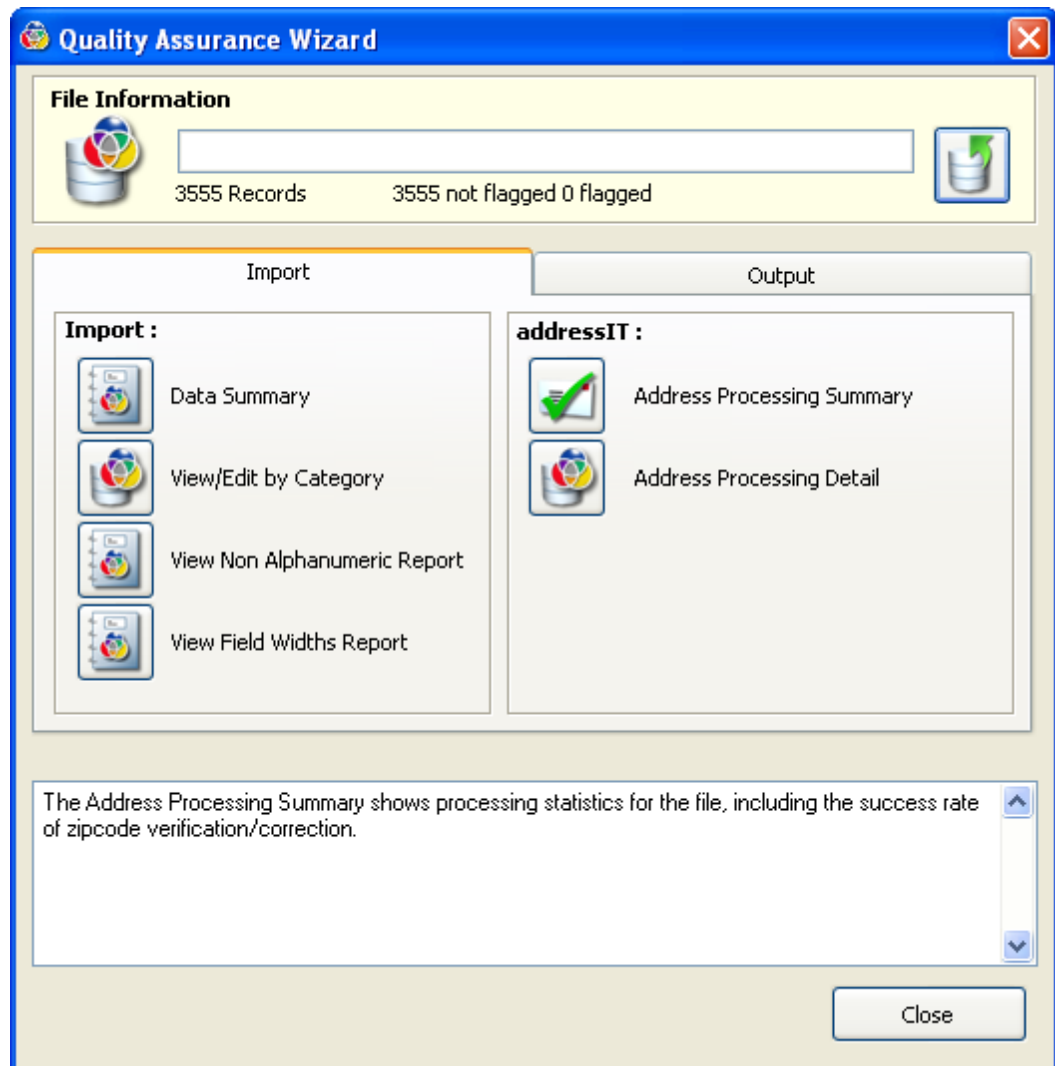
Data Summary

This is a summary of your import data. It summarizes information such as:

- Potential data errors

- Data enhancement and relocation

- Records which should perhaps be excluded from the output

- Salutations

- Male/Female split

- Forenames found in matchIT's names table.

addressIT Summary

The addressIT Summary contains zip coding statistics for the Main File, including the results of zip code generation and correction.

View Records by Category

This allows you to display all records meeting selected categories, most of them as categorized on the Data Summary which you can view after Import.

You can select several categories to view at once. You can also choose to delete the selected records, print them and/or to write them out to a separate file.

After you browse the selected records, matchIT returns you to the same screen to select more categories if you wish. To leave the screen, select **Close**.

View Data

View Data allows you to browse and update the selected Main File in any order.

Click **Continue** to go on to the next stage of the Quality Assurance Wizard or Finish to exit the Quality Assurance Wizard.

Quality Assurance Wizard – Output
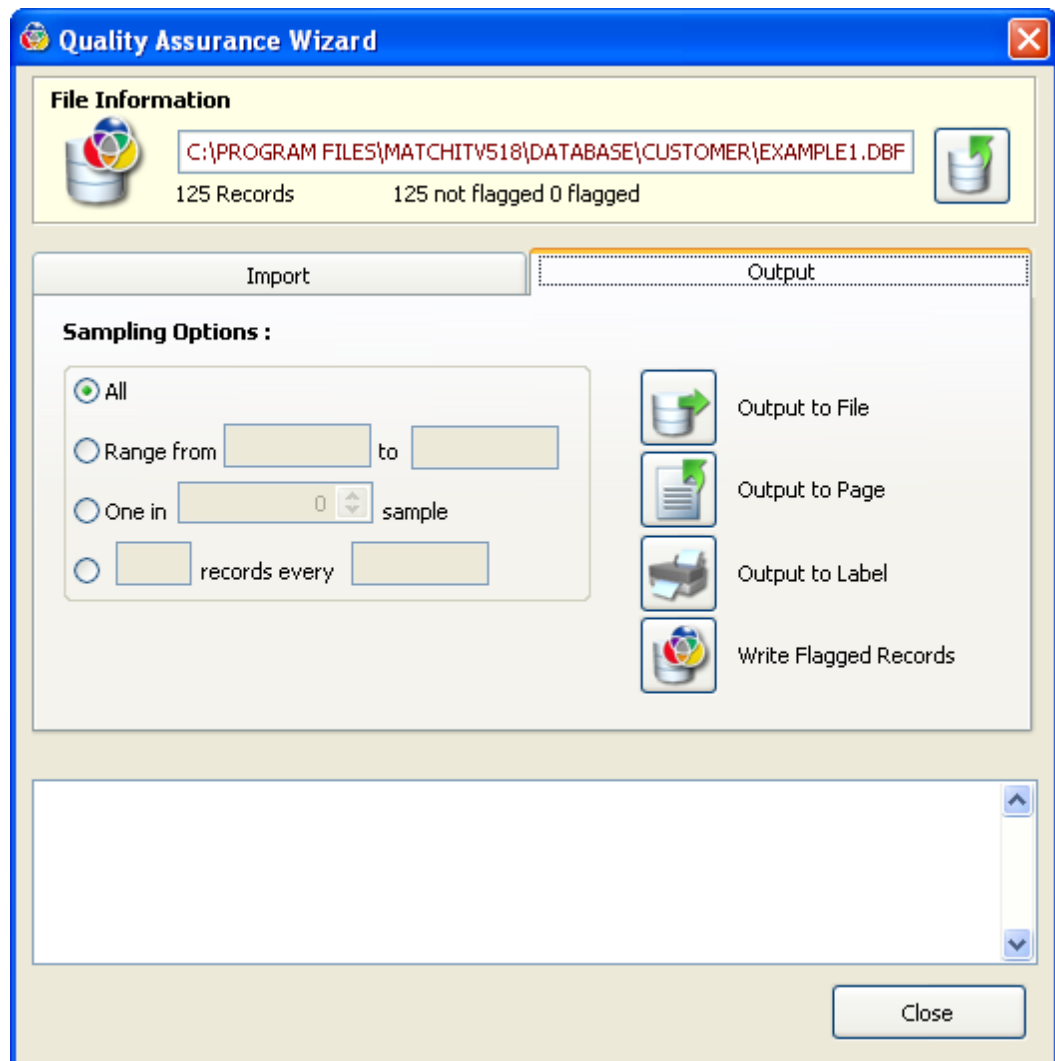
All

With this option you can choose to output all records.

Range from

With this option you can choose to output all records or a selection of records.

One-in-N Sampling

When you choose this option, an intermediate table is created containing a selection of the records in your Main File. For example, if you have a Main File of 100,000 records, and choose One in 1000 sampling, your test file will contain 100 records – the 1000th, 2000th, etc. from your Main File. Obviously, sampling does not guarantee that there will be no errors in your data.

Output to File/Output to Page/Output to Label

These buttons will bring you to the standard matchIT Output screen, from which you can access all of the output options available in your version of matchIT, but for the selected sample of records. Even if you never produce labels from matchIT, using the default Address label to preview a sample of names and addresses gives a clearer picture than a Browse view of what those names and addresses will look like if and when they are eventually printed from another piece of software. You can even print a sample of "labels" on plain paper to Q/A them more easily and perhaps file them.

Write Flagged Records

matchIT will now copy the flagged records from your Main File into the new File you just named and unflag them from this new file. When it has finished, matchIT will automatically select the new file as the Main File.

View Data

View Data allows you to browse and update the selected Main File in any order.

Click **Close** to exit the Quality Assurance Wizard.

Quality Scoring

Address Quality

There are two levels of address quality reported; the Pre-Clean quality is reported prior to any cleaning or reformatting – the idea behind this is to grade the quality of a users data before cleansing occurs so that the uplift can be compared.

The quality score will be assigned only if there is a field labeled QUAL_ADDR (char 2) in the Main File Layout. This field can be added in the Setup Wizard.

A separate flag denotes if the record is foreign (non-US).

We will provide a quality score for address on a scale of 0-9, with quality scores being assigned by as follows:

- 0 = sdf i.e. garbage strings identifiable by pattern recognition

- 0 = empty

- 1 = empty address key

- 3 = no city detected and no zip code entered

- 4 = no street detected

- 5 = no city detected and zip code entered

- 9 = Maximum score (one point for each region and country entered / one point for premise detected / two points for each of address, town and zip code entered)

The quality scores are assessed in the order shown and as soon as a quality condition is met, that score is assigned and the scoring process ends.

Parameterized whether to score points for country, region or premise using flags in Params.dbf as follows:

- QS_REGION – 0 disables scoring for County

- QS_PREMISE – 0 disables scoring for Premise

- QS_COUNTRY – 0 disables scoring for Country

QS_BLANKPC – 0 restricts addresses without postcodes to a score of 1.

With the settings above, the scores in the table below should be returned, Pre-Clean:

Quality	Description
7	address line 1, town, county, zip code populated.
6	address line 1, town, zip code populated.
5	address line 1, zip code populated.

Quality	Description
4	Town, zip code populated
0	empty address and zip code
1	none of the conditions above

Name Quality

The quality score will be assigned only if there is a field QUAL_INDIV (char 2) in the Main File Layout. The user can add this field in the Setup Wizard.

We will provide a quality score for name on a scale of 0-9 (with 0 being the poorest quality), with quality scores being assigned as follows:

- 0 = Empty
- 0 = sdf i.e. repeated patterns denoting garbage
- 0 = Empty name key
- 1 = John
- 2 = Smith
- 3 = J Smith
- 4 = Mr Smith
- 5 = Mr J Smith
- 6 = Joh Smith (Unrecognized first name or error in first name, no prefix)
- 7 = John Smith
- 8 = John R Smith
- 9 = Mr John Smith
- 10 = Mr John R Smith

The quality scores are assessed in the order shown and as soon as a quality condition is met, that score is assigned and the scoring process ends.

To generalize the examples above, the scores in the table below should be returned, after any name parsing, casing or salutation generation:

Quality	Description
10	Surname, input title, forename and middle initial(s) all populated
9	Surname, input title, forename all populated, no middle initial
8	Surname and forename and middle initial populated but no input title
7	Surname and forename recognized in Names.dat populated but no input title
6	Surname, and forename unrecognized in Names.dat, populated
5	Surname, title and initials populated, no forename or length of forename is 1 character
4	Surname and title populated, no forename or initials
3	Surname and initial(s) populated but no title or forename
2	Surname populated but no title, initials or forename
1	Empty surname

Email Quality

The quality score will be assigned only if there is a field QUAL_EMAIL (char 2) in the Main File Layout. The user can add this field in the Setup Wizard.

We will provide a quality score for email on a scale of 0-9 (with 0 being the poorest quality), with quality scores being assigned as follows:

- 0 = Empty

- 0 = Nonsense

- 9 = Username, domain, and recognized top-level domain (TLD) found

- 2 = Top level domain (TLD) not recognized

Campaign Processing

This section is only applicable if you have matchIT **Campaign**.

Introduction to Campaign Processing

The Campaign module of matchIT allows you to select records for mailing campaigns and keep track of who was selected for each campaign. Each time you produce a Mailshot or output file, matchIT will automatically update your Campaign History table with details of the current campaign. Later on, you can examine this history table.

To maintain Campaign History automatically, you will need to add the fields **MAILING_ID**(width 10) and **MAIL_DATE**(Date type) to your Main File Layout. The Setup Wizard will normally add these fields automatically, but you can add them manually by specifying them at the end of the Setup Wizard or via Main File Layout option in the Tools menu.

You will also need a History table. Whenever you update mailing history, matchIT will ask you for a History table. You can create a new History table at any time by entering a new name when asked. We recommend that you use separate folders for each campaign that you are tracking, so that the history tables for each campaign are kept separate.

The Campaign Process

There are 5 main steps involved in the campaign process in matchIT:

- Firstly, import and deduplicate your data in the normal way for matchIT. Basically, you should create the table which (or part of which) will be used in the output step.

- Secondly, run a query if you want to select a subset of the records on the table for output. You may also use a separate file of reference numbers (such as one generated by analysis software) to filter the output.

- The next step is to create the actual output. You may need to create a Campaign record to record the Campaign Identifier and date. When the output is created, the actual updating of the history file usually takes place.

- Some users producing multiple mailshots or output files from different selections of the master file may find it more efficient to create all the output first, and then update the mailing history file afterwards. In this case, an additional step (updating the history file) will be needed.

- Viewing Campaign History allows you to see when each record in the Main File was selected, for which campaigns.

Campaign Options

The Campaign Options can be accessed through the Output screen's Campaign Options tab or by running a Query with Use Campaign History Options switched on.

- **Campaign ID** – Either select one from the list or add a new ID by clicking the Add button. Campaign IDs can be deleted by clicking on the Remove button.

- **Date** – Enter the date the mailshot is to be sent.

- **Use Reference File** – Switch this option on to use an external list of references to filter the output/query. The list of references should be in a comma-separated ASCII file format (basically, one reference per line). The reference length should be the same as the length of UNIQUE_REF in your Main File. For each reference that matches to a UNIQUE_REF in the Main File, one record will be output. Click on the *Choose Another* button to select the reference file. matchIT will remember the file name the next time.

Campaign File Splitting Options

The "Split file for a Campaign" option is located under the Campaign Processing section of the Output menu.
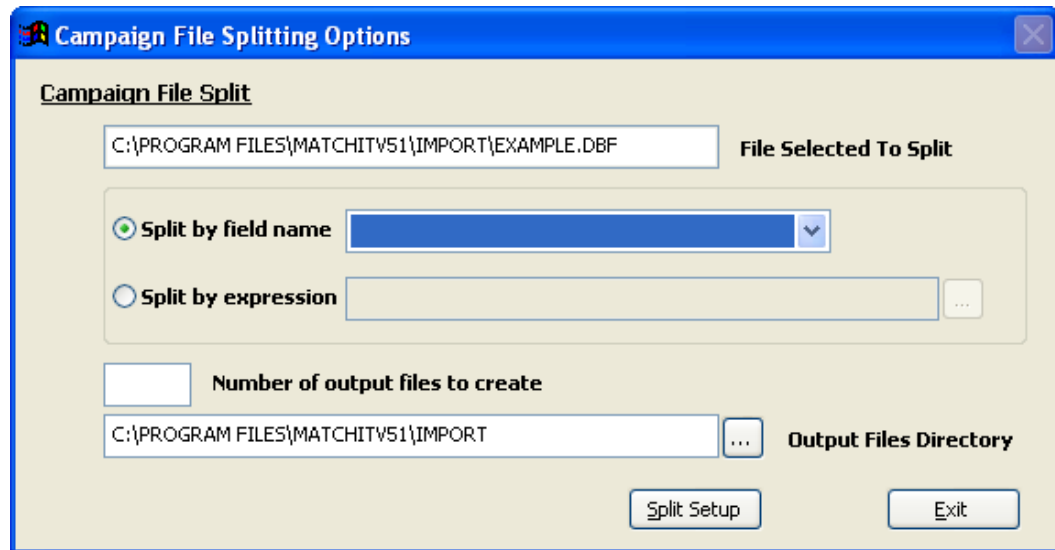
The above window allows you to control what file to output and what field/expression to split on. You can also specify how many output fields you would like created, and where they need to be stored.

Campaign History

Creating and Updating Campaign History

To create or update a Campaign History table, all you need to do is Run a Query or go through the steps to produce output with Output to File or Printer. In either case, you must switch Use Campaign History Options on and set those options accordingly (See Campaign Options). It should be noted that if you produce an output as a print preview, this will NOT cause Campaign History to be updated. Also, you must have the fields MAILING_ID and MAIL_DATE in your Main File to trigger matchIT's Campaign History function.

When you produce a mailshot, either by running a Query or by Generating Output, matchIT will ask you if you would like to update a Campaign History file.

If you answer 'Yes' to this question, you will then need to choose a History table to be updated. If you have not already created a history for this Main File, just type in the name of the new history table in the file selector. matchIT will automatically create a new History table if it does not exist.

If you answer 'No' to the above question, updates can always be made later.

Viewing Campaign History

To view Campaign History, choose the **View Campaign History** menu option from the **Output** menu. matchIT will now ask you for a History table. You should choose your Campaign History table for this Main File.

Queries

Query Overview

Queries provide the facility to search a chosen table for all records whose contents have characteristics that you specify. You can define standard queries that you can repeat later or modify.

Queries are manipulated through the Query Manager.

Query Manager

Using the Query Manager, you can create a query or run an existing query.

Query Name

The name of the current Query will be displayed if you have either Saved the Query or Opened an existing Query.

Filter

Filter Conditions are used to limit the records that will be included in the Results Table. For example, a filter of COUNTRY#'US' would output any records that did not have a Country value equal to "US".

Fields

The Fields page is used to select the fields that will be included in the Results Table.

Order By

On the Order By page, select the fields by which you want the Results File ordered.

User-Defined Clauses

Additional SQL Select clauses can be added in the User-Defined Clauses page.

Use Campaign History Options

Check this to use Campaign History Options. Your Main File must contain the fields MAIL_DATE and MAILING_ID. This option cannot be used with a Query containing a GROUP BY clause in the User-Defined Clauses. When you run a query with this option switched on, you will be brought to the Campaign Options screen.

View SQL

To edit the SQL Select statement directly, click on the View SQL button.

Clear All

Click on the Clear All button to remove all of the filter conditions, field selections and clauses you have entered.

New Query

Clicking on the New Query button will close the existing Query without saving changes.

Run

After either creating a new Query or Opening a previously saved Query, click on the Run button. Alternatively, select Run Query from the Query menu and select a previously saved Query. You will be prompted to name the Results File. This is the file in which the results of your Query will be stored.

Open

To open a pre-existing Query, click on the Open button.

Save

To save the current Query, click on the Save button. You will be prompted to name a Query File.
 Once saved, a Query can be opened in the Query Manager or run by selecting Run Query from
the Query Menu.

Query Manager – Filter
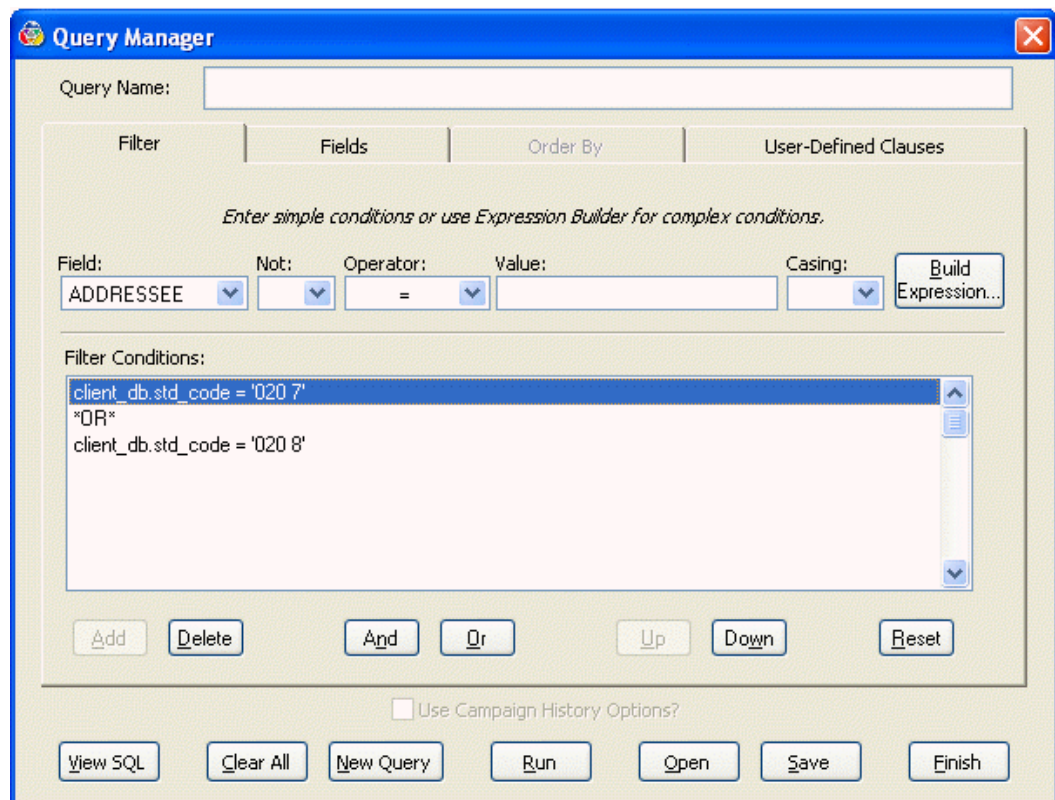
To enter a simple condition, choose the Field and Operator, enter a Value and choose a Casing
option (if desired), then click on the Add button. To build a more complex condition, click on the
Build Expression button.

Field

This drop-down list contains all of the fields that exist in the Main File. Choose the field for
which you want to create a condition.

Not

Select Not to negate the condition.

Operator

Choose an operator for the condition.

Value

Enter the value for which you are testing. When the Empty Operator is chosen, Value will be unavailable.

Casing

The Casing option selected will be applied to both the Field and Value you have selected. If you have chosen Upper, then the Uppercase of the field will be compared with the Uppercase of the value you have entered. When the Empty, Length > and Length < Operators are chosen, Casing will be unavailable.

Build Expression

Click this button to access the Expression Builder. This will allow you to build more complex filter conditions.

Filter Conditions

This window displays the filter conditions that will be applied to the query.

Add

Click the Add button to add a condition to the Filter Conditions window. The Filter Condition added will be based on the values of Fields, Operator, Value and Casing.

Delete

Click on Delete to remove the currently selected condition from the Filter Conditions window.

And/Or

To create multiple condition in your Query, click on the And/Or buttons to add the appropriate connector to the Filter Expressions window.

Up/Down

Click on the Up/Down buttons to rearrange the filter conditions in the Filter Expressions window.

Reset

The reset button will remove all of the Filter Conditions you have entered.

Query Manager – Fields
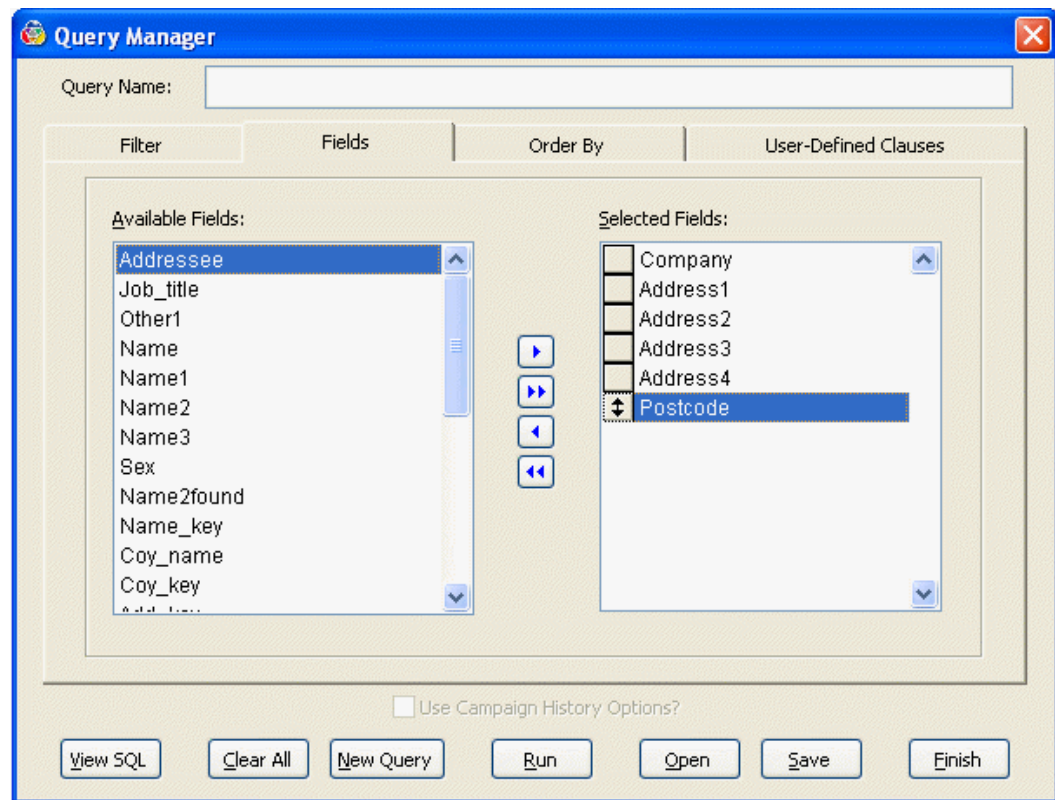
Available Fields

This window will display the fields that exist in the Main File. To transfer a field from this window to the Selected Fields window either click on the single right arrow button or double click on the field name itself. Clicking on the double right arrow button will move all fields from the Available Fields window to the Selected Fields window.

Selected Fields

This window will display the fields that will be included in the Results Table. To remove a field from this window either click on the single left arrow button or double click on the field name itself. The field will then be moved back to the Available Fields window. Clicking on the double right arrow button will move all fields from the Selected Fields window to the Available Fields window.

NOTE: If no fields appear in the Selected Fields window, all fields that exist in the currently selected table will be included in the Results Table.

Query Manager – Order By
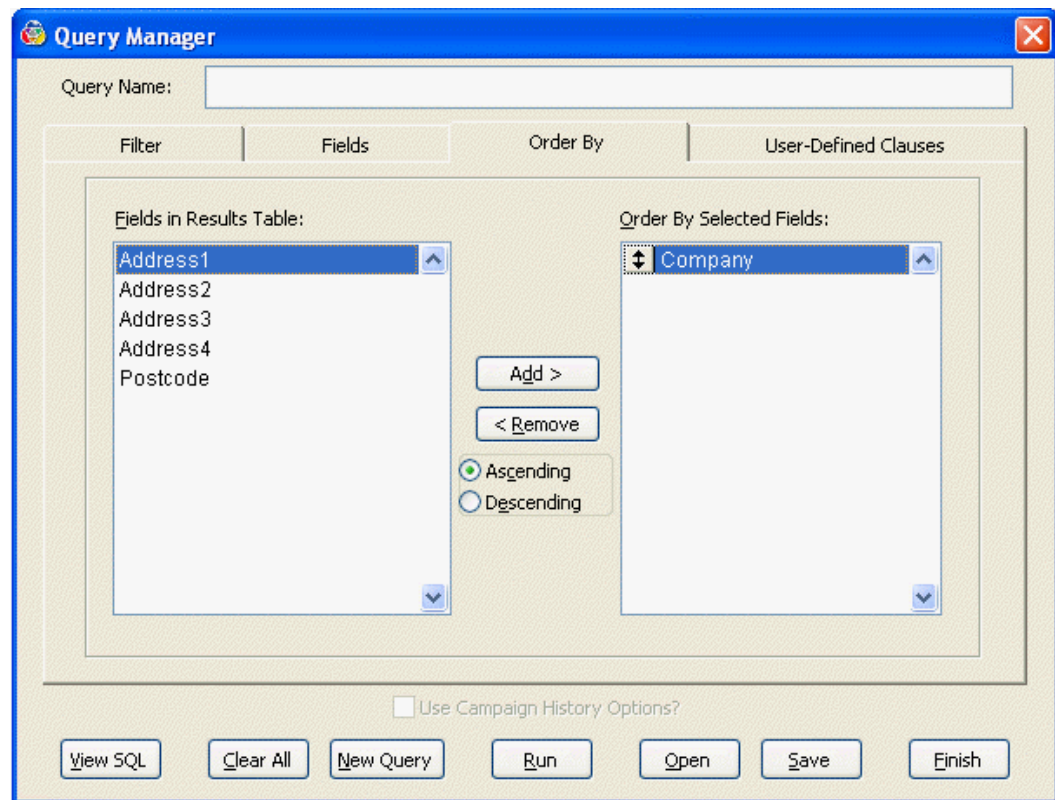
Fields in Results Table

This window will display all of the fields you have selected to appear in the Results File. To transfer a field from this window to the Order By Selected Fields window choose Ascending or Descending, then either click on the Add Button or double click on the field name itself.

Order By Selected Fields

The Results Table will be ordered by the fields listed in this window. To remove a field from this window, either click on the Remove button or double click on the field name itself.

Query Manager – User-Defined Clauses
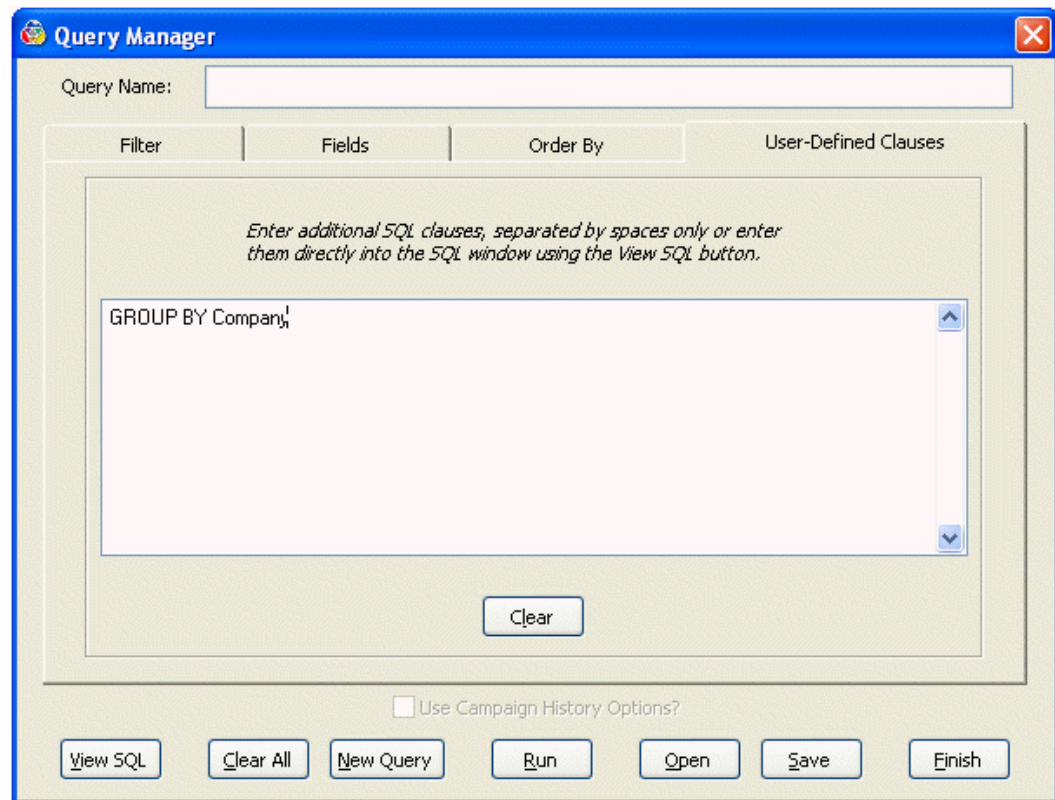
Any additional SQL Select clauses you wish to use can be entered in the User-Defined Clauses window. Alternatively, you can edit the SQL Select statement directly using the View SQL button.

Clear

Click on the Clear button to remove all User-Defined Clauses you have entered.

Count Values

Select Count Values from the Query menu to create a simple query to count the occurrences of the values contained in a field (or field expression) of your Main File. You can Run this query from the Count Values screen or Save the query, which can then be opened in the Query Manager.

Run Query

Select Run Query from the Query menu to run a previously saved Query. If the Main file and the Query comply with the Campaign requirements, you will be asked if you want to Use Campaign History Options. If you answer Yes, you will be brought to the Campaign Options screen. If you answer No, or upon exiting the Campaign Options screen, you will be prompted for a Results File name. This will be the name of the file containing the results of your Query.

System Commands Overview

If you know some FoxPro (matchIT uses Visual FoxPro) or dBase (known as "xBASE") programming, you can use the System Commands utility to key in simple xBASE commands. The System Commands facility is not intended to replace any requirement for a full FoxPro license rather than the run-time one included with matchIT, and your support provider does not provide telephone support, Online Help or documentation for program commands other than the examples below.

As all files are closed when you select the System Commands option from the Tools menu, you will have to first use the database you want, or select it after keying any command that requires a database.

Some FoxPro commands that you might find useful are (type the command in the box then click "Process Command"):

- To open a database file for use by other commands:

  `USE ?`

  (then select the database)

- To set the order of a database:

  `SET ORDER TO UNIQUE_REF` (if the index exists)

  `INDEX ON ADDRESS1 TO TEMP` (if the index doesn't exist)

- To browse (view) a database:

  `BROWSE`

  (This allows various facilities to reposition on a different record, seek a key etc. through the Browse menu)

- To delete all records from a database from the current record onwards:

  `DELETE REST`

  In this example, `REST` is a scope clause - scope clauses can be added to all the commands below. Other examples of scope clauses are `ALL`, `NEXT 100`, `NEXT 9999` etc.

- To recall logically deleted records from a database for a condition:

  `RECALL FOR` condition

- To delete all records from a database which have blank addressee AND company fields, OR blank address lines 1, 2 and 3:

  `DELETE ALL FOR (EMPTY(ADDRESSEE) AND EMPTY(COMPANY)) OR`
  `(EMPTY(ADDRESS1) AND EMPTY(ADDRESS2) AND EMPTY(ADDRESS3))`

  In this example, `FOR` specifies a condition clause - you can combine condition clauses with `AND` and `OR` and use brackets to control the way in which the conditions are combined. Note the additional brackets in this example, surrounding the first condition (to the left of the `OR`) and the second condition (to the right of the `OR`)

- To initialize a field called `LISTSOURCE` to a value `'ABC'` if it is blank, for a given range of unique reference numbers:

```
REPLACE ALL LISTSOURCE WITH 'ABC' FOR EMPTY(LISTSOURCE) AND
BETWEEN(UNIQUE_REF,  '0001000',  '0099999')
```

- To copy records for various list sources to another file and then delete them from the first file:

```
COPY TO  drive:path\filename  FOR INLIST(LISTSOURCE, 'ABC',
'XYZ', '123' etc.)
```

```
DELETE FOR INLIST(LISTSOURCE, 'ABC', 'XYZ', '123' etc.)
```

- To create a copy of the file in a sorted order:

SORT TO drive:path\filename ON fieldname1, fieldname2, fieldname3 etc. FOR condition

Configuring Centralized Files

Configuring Centralized Files

It is recommended that matchIT be installed to the Centralized File Server (CFS) to provide easy access to the configuration of the Centralized Files. The CFS can be any Windows based hardware capable of running the matchIT Suite.

Install matchIT to the CFS or select an existing matchIT client as the CFS. Ensure all Suppression options are configured correctly and function normally.

If you choose to use an existing matchIT client as the CFS make sure it is NOT using the Centralized files as its default Suppression data. Copy the data to a new folder e.g. C:\CSuppression or C:\CDAT, where C indicates Centralized.

Centralized Suppression files

On the machine selected as the Centralized File Server (CFS):

1. Install/copy your suppression files to e.g. C:\CSUPPRESSION

2. Copy the file SUPPLOCS.DBF (from the \matchITV51 directory) to the CFS dir e.g. C:\CSUPPRESSION\SUPPLOCS.DBF

3. Share this folder (allow all users, read only access) to enable access from the network via mapped drive or network share

On the matchIT clients (MIC):

1. Install/copy your suppression files to a local disk and set suppression file locations on the local copy of matchIT as usual (menu SUPPRESSION>SUPPRESSION OPTIONS> STANDARD SUPPRESSION OPTIONS). From within matchIT (using the menu VIEW>VIEW TABLE and change 'Files of type' to 'All files'), open MATCHIT.INI and under the CSUPPLOCS field, enter the network path to your CFS where the SUPPLOCS.DBF file resides e.g. \\CFS\SUPPRESSION , then close MATCHIT.INI.

2. Open SUPPLOCS.DBF on the CFS e.g. \\CFS\CSUPPRESSION\SUPPLOCS.DBF.

3. Edit the path for all files from e.g. C:\CSUPPRESSION\DISCONNECT\DECEASED.DBF to reflect the network path e.g. \\CFS\CSUPPRESSION\DISCONNECT\DECEASED.DBF

When matchIT is started on the MIC a check is done comparing the local suppression files (as specified in the local SUPPLOCS.DBF) to the centralized suppression files. If the local file is older, the user is prompted to update the local copy. If they reply yes, ALL out of date MIC files are updated with CFS files. If they reply no, a warning is displayed stating that no local suppression files will be refreshed. Due to the size of the Suppression files, copying the entire set may take up to 10 minutes depending on network speed/traffic.

Centralized DAT files

On the machine selected as the Centralized File Server (CFS):

1. Copy matchIT DAT files to C:\CDAT

2. Share this folder (allow all users, read only access) to enable access from the network via mapped drive or network share

On the matchIT client (MIC):

From within matchIT (using the menu VIEW>VIEW TABLE and change 'Files of type' to 'All files'), open MATCHIT.INI and under the CDATLOCS field, enter the network path to your CFS where the CDAT files reside e.g. \\CFS\CDAT, then close MATCHIT.INI.

When matchIT is started, ALL files in the CDAT location with file extension DAT and also Business noise words.DBF and mail_exclusions.dbf will be copied invariably to the local DAT location in PARAMS.DBF.

The DAT location may also be viewed/edited in matchIT via menu: JOB/SETUP>OPTIONS>FILE LOCATIONS, DAT FILES PATH

NOTE: The matchIT DAT location must be within the matchIT directory 'C:\Program Files\matchITV51' for update to take place. This restriction enables a user to use a non-standard local set of DAT files for a specific job (via matchIT Options, File Locations). Customized DAT files stored outside of the matchIT directory will not be updated.

Database Utilities

Functions of Database Utilities

You can run Database Utilities from the matchIT Program Group, or run FIXDB.EXE in the matchIT directory. This utility can be used for many different things, for example:

- modify layouts of matchIT's parameter tables

- search for corrupt tables or indexes in the matchIT directory tree

- rebuild corrupt indexes

- correct a corrupt database header

- permanently delete "logically deleted" records

- recall "logically deleted" records

- wipe a whole table permanently of all records

- view and report on any database (table)

- change a delimiter in an ASCII text file

- append to any table from DBF and text files

- export from any table to DBF and text files

- regenerate matchIT's UNIQUE_REF so it is unique

- blank out a specific field in a database

- eliminate blank address lines by shuffling up later lines

- limit field widths to the maximum size of the data within them

- generate an Addressee field from Prefix, Forenames and Surname.

Only the most commonly used options are described below; the remainder are described in the Online Help.

Use the normal matchIT menus in preference to Database Utilities whenever possible as Database Utilities does not perform the usual safety checks that matchIT does.

File Menu Options

Choose Database: to open a different database (table) from that selected previously

Convert Text File: while matchIT can cope with most standard data formats, occasionally you may have to deal with a non-standard data file. If you want to import a delimited file which has no Carriage Return Line Feed (CRLF) at the end of each record, matchIT has to convert it before Import. This option will convert such a file into comma delimited format (CSV). You will need to know how many fields there are in your data file and what the delimiting character is. Follow the on-screen prompts. You will then be asked for a filename for the newly created file.

Append: appends records into a DBF file from another DBF file, comma or tab delimited file, or an SDF (fixed length) file. This can be useful when you want to create a DBF quickly without matching, perhaps to do some reformatting before asking matchIT to process it. The field names and file layout must be the same before appending.

Export Database: exports a file to a choice of formats, with a new file name.

Quit Database Utilities: to return to Windows.

Database Menu Options

All of these options ask you to select a database if one is not in use.

Modify Structure

This option allows you to change the structure of any database (change the fields in a table).

Undelete

If you select the Undelete option, all logically flagged records are recalled to the database, e.g. from when you flagged matches.

Pack

Permanently deletes records from the table.

Zap

Permanently deletes all records (not just logically deleted ones).

Fix Header

This corrects a DBF file header where the number of records in the header is wrong, e.g. if you ran out if disk space during an Import

Recreate Index

A DBF file cannot be used if it has a corrupt CDX index attached to it. This option will remove the index and then create a new one. It is the equivalent of deleting the CDX file through Windows Explorer and then using Reindex to index the DBF.

Reindex

Recreates indexes for a DBF file (including matchIT's files, not just your Main Files).

Browse Last/Reset Browse

These options allow you to change data, recall deleted records etc. In Database Utilities, unlike in View Table in matchIT, deleted records are visible but are denoted by a thick black line in the left margin. The Reset Browse option also makes the Browse window display revert to the default configuration, allowing you to set up a new view if you have a problem with the current view.

Count Records

This option allows you to count the records in the database from the record that you are positioned on (for example, after Browsing) to the end of the file. If you have just opened the file, this will be the total number of records in the table.

Field Menu Options

Generate Unique Refs

Creates new unique references in the file, starting at the number you enter. This ensures the reference numbers are unique, for example if you have been having problems with matching. After selecting this option, you will need to redo any matching performed on the File.

Check Unique Refs

Checks all the reference numbers (in a field called UNIQUE_REF) are actually unique. If they are not, use generate unique references to recreate them (see above).

Shuffle Up Address Lines

If there are gaps in your address lines, either because of the original database layout or because matchIT has extracted, say, company names and zip codes, you can remove them using this option. It will move all address lines to the 'top' by filling Address1 first, then Address2, etc.

Change Fields to Fit

Changes the width of all character fields in the Main File to the maximum width of the data in each field. This does not change or empty any of matchIT's key data fields.

Generate Addressee

Creates a freeform name field from Prefix, Forenames and Surname – use this when the contents of Prefix, Forenames and Surname are not reliably the components of an individual's name e.g. Prefix = "The", Forenames = "Managing" and Surname = "Director".

Report Menu Options

List Structure

This option allows you to list (i.e. print out) the structure of any database.

Check Files

This facility goes through the matchIT directory and all its sub-directories, attempting to open all of the database it finds. If there is an error opening a databases, this is reported to the screen, along with the name of the database, and the reason for the error. This can be extremely useful when there has been a hard disk crash, for instance, and a database (or index) has become corrupted. This option will be able to track it down straight away.

Problems

Error Log

When an error occurs in matchIT, it is recorded in a log, along with some useful diagnostic information. You can examine the Error Log from the View menu. matchIT positions on the last error to occur – scroll up to see previous entries. Close this screen by pressing Esc or clicking the X in the top right corner. matchIT will then ask if you wish to print any of the entries – it will start with the last one.

When calling the support line, you may be asked to look up the program and line number recorded in this log. Also, not all errors are displayed in a message box on the screen – often the cause of the problem occurred earlier in operation. You should always scroll back in the error log, looking for the first record with the appropriate date and time.

Print Diagnostic Information
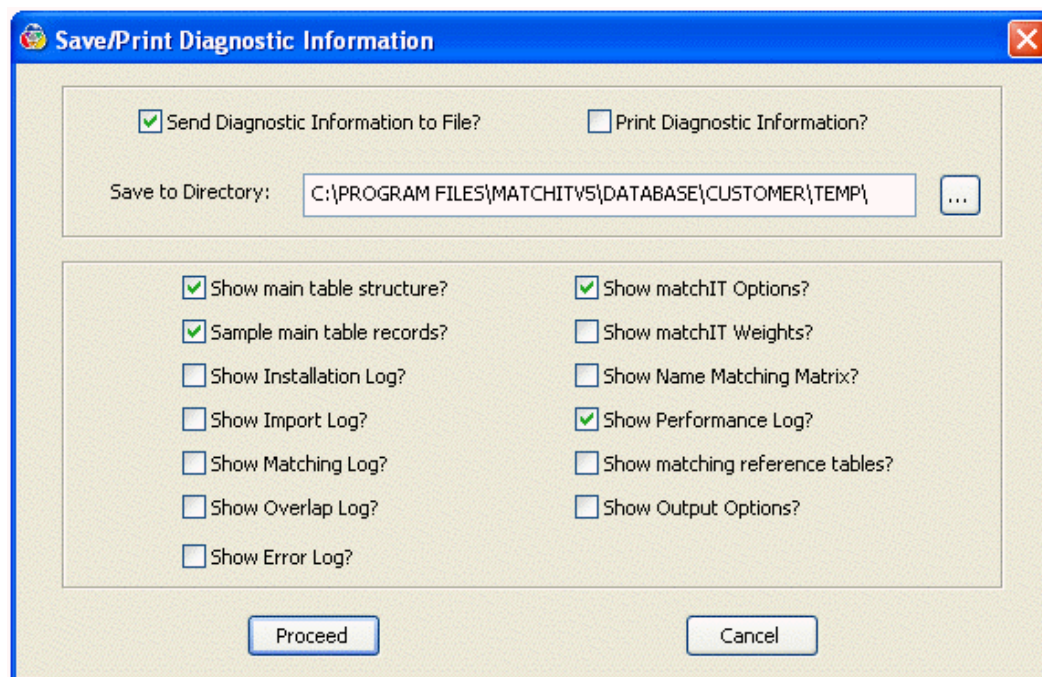
From the Tools menu, this option allows you to easily send diagnostic information to a printer or file. If you send the report to file, it is advisable to send them to a clean directory – it may get a bit confusing otherwise. This allows you to simply see your current matching setup, the layout of your Main File, sample records and/or the logs matchIT produces. For example, these can help you work out why a matching run did not produce the expected results. These reports are:

- Main File Structure – the layout of the selected Main File

- Sample of Records – a dump of the first 100 and last 20 records in the selected Main File

- Installation Log – the date and time the installation started, and the files that were copied or replaced.

- Import Log – a record of the last import. If an import was aborted for whatever reason, the last record number will be shown – otherwise it will contain the number of records in the database.

- Matching Log and Overlap Log – The number of records in the database and the number of comparisons made while matching.

- Error Log – the contents of the MIERROR.DBF.

- matchIT Options – the contents of the PARAMS.DBF.

- matchIT Weights – the contents of the WEIGHTS.DBF.

- Name Matching Matrix – the contents of the NAMEPARM.DBF.

- Performance Log – the contents of the PERFORM.DBF.

- Matching Reference Table – the contents of the MATCHES.MDF.

- Output Options

Solving Problems

Error Messages

matchIT deals with errors by referring to a database of FoxPro error codes (MIACTION.DBF). it then takes the action indicated by the entry, in the database for that error. Full information about the error is recorded in the MIERROR database, this information can be vital for your support provider or your own PC support people to diagnose the problem. Each record in the MIERROR database is very large, so if the MIERROR.DBF grows to more than 100 records it is automatically cut back to the last 50 errors logged. The third database involved is MIPERSON.DBF. This contains two records; the first record contains the title, name and extension number of the on-site person responsible for supporting matchIT: you may change the details here by browsing the database and overkeying the information. The second contains details pertaining to your support provider, for reporting problems which may not be able to be resolved on-site.

If the error which occurs does not have a corresponding entry in MIACTION.DBF, the default error 99999 is used. You can update or add entries to MIACTION.DBF, to specify different treatment of the error. However, this should only be done on the advice of your support provider.

Accordingly, a list of error messages, their meaning and the action you should take is not printed here: in the event of an error, the relevant information will be displayed.

When the error display says "We'll carry on from where we can", matchIT waits a few seconds for a response, then carries on with the next step of the program. If you want to abort the process and return to the menu, press Escape.

Problems Caused by Power Failure Etc.

Power failure or hardware/system software "glitches" causing data corruption are a common source of problems. This is usually during Import or Matching, as they are the longest duration processes. Any database the system is writing to at the time could be corrupted, giving the message

"NOT A TABLE"

when it is next used. The next line will contain the name of the database causing the problem e.g. MATCHES.DBF in the example in the previous section. If so, restore this file (which may be in one of the sub-directories of matchIT) from backup.

If you are not sure which file is corrupt, or have some problem starting matchIT, you can use the Check Files (on page 165) option in Database Utilities to find out which file(s) report problems.

You can selectively reinstall the versions of any files originally delivered: to do this, run the SETUP program on the installation CD (as for the initial installation) but add /X to the SETUP command e.g. run

D:SETUP /X

You will see a list of the installation files displayed on the screen. They are listed alphabetically by type i.e. all the DBF files are grouped together and all the FRX files are together etc. You need to select the file(s) you want to restore, and the directory you want to restore them to.

If the problem occurred during the Import run, the most likely database to be corrupted is the selected Main File. If you have modified the layout, there will be an earlier backup file in the DATABASE directory with the same name but with a file extension of BAK instead of DBF: rename it (in Windows Explorer) as *.DBF; if there is no such file, copy the Main File from backup.

If the problem occurred during matching, the most likely databases to be corrupted are MATCHES.MDF, MERGES.MDF in the Main File directory, PERFORM.DBF in the DATABASE directory, or MATCHES2.DBF and/or SETS.DBF in the MATCHIT or REPORT directories (although any other database could have been corrupted, depending on the operation being performed). Copy files as necessary from backup, or the matchIT install disk. matchIT databases also require index files with an extension of CDX. In some circumstances, databases can be okay but index files corrupted. If a CDX is missing (or corrupt) matchIT will report the error.

"INDEX DOES NOT MATCH DATABASE FILE. RECREATE INDEX"

and/or

"DATABASE IS NOT ORDERED"

Index files can be recreated by using the Database Utilities Recreate Index option. Database Utilities can be entered by selecting the appropriate icon in the matchIT program Group in Windows, or by selecting the Database Utilities option from the matchIT Tools menu.

If you can't load matchIT at all, try entering Database Utilities (from the program group). Then choose Check Files, from the Report menu.

When you have recovered all the corrupted files, you will need to start the run again from the beginning of the step at which the problem occurred.

Matching Problems

Overview of Matching Problems

When reporting matches or overlap between files, matchIT looks up the full details for each record involved by using its unique reference number as stored in the Matches or Merges database.

When investigating problems with matching, before doing anything else, browse the Main File and check that the imported data looks okay, including the derived fields such as NAME, NAME1

etc. The Key fields are fairly obscure, but they should not in general be blank. If they are, or if e.g. the NAME2 field is blank, it could be because you have not allocated the right field names to the data (or the Intelligent Setup Wizard has not correctly identified the fields). You should check in particular that the right names have been used for people's names – usually ADDRESSEE for a name keyed all in one field and PREFIX, FORENAMES (or INITIALS), SURNAME when they are split up.

Too Few Matches or Scores Too Low

Possible causes of these problems are:

- the Minimum Score to Report is too high

- the Weights are poor, or are placed on un-normalized fields such as Addressee

- the primary Match Keys you have used have caused matchIT not to look at some potential duplicates e.g. you have used NAME_KEY + ZIP in one step, instead of ZIP in one step and NAME_KEY or NAME_KEY + LEFT(ZIP,5) in another step, or you have used an un-normalized field as a primary match key, such as Addressee

- you may have Match on Location set on in "Matching Setup" (on page 97) when you don't want to use it. If you want to consider matches based on e.g. contact and company names irrespective of location (address and zip code), you should uncheck this box.

- although you did more than one Find pass through the data, with different keys, you said "Yes" to "Is this a new Analysis" on the second pass, when you should have said "No."

If you still can't resolve the problem, find two records which should be reported as duplicates and contact your support team or your support provider.

Too Many Matches or Scores too High

Some of the checks above are worth looking at, in case the problem is the reverse of that described above. You may also benefit from using primary Match Keys that do not allow some of the false matches to be reported (e.g. where records are being reported because they match on name and have a blank zip code, but the addresses are different, use LEFT(ZIP,5) + NAME1 starting at zip codes beginning with 'A', instead of NAME1 + LEFT(ZIP,5).

If you are dealing with foreign data (not from an English-speaking country), you may need to add common words to the Names and Words table to stop matchIT from paying any or too much attention to them when comparing records e.g. add Weg as an **Address** word for Holland and the Scandinavian countries.

If this does not solve the problem, see the section below.

Totally False Matches Being Reported

Possible causes of this are:

(a) Unique References are not Unique

If when you merge two databases and then View Matches or View Overlap, records are displayed with no similarities whatsoever, it is possible that the records do not have unique entries in the UNIQUE_REF field. You can check the unique references by using Database Utilities, Check Unique Refs. If this field was not allocated by matchIT on Import, rename your input field as

some other name (URN will do) and define a new field for UNIQUE_REF and matchIT will allocate references that are unique.

When you import two databases for merging, be sure to start the next reference number in the second database higher than the last record number in the first database e.g. if the unique reference number of the last record in the first database is 5000, the next record number of the second database needs to be greater than 5000.

If you have merged databases together and forgotten to give them unique references when originally importing them, you can use the Database Utilities option [Generate Unique Refs](#) to regenerate them. You must then Find Matches or Find Overlap again to find the matches for the new reference numbers.

If this happened in a Job Script, make sure that you set the Next Reference Number in the Options (Input Options) for each step, so that each database has a unique range of references, or use the Multiple File Wizard which ensures uniqueness.

(b) Inappropriate Report

Alternately, the problem could be due to having selected the wrong report. When choosing a report, make sure you select one appropriate to the options you have defined. If using one of matchIT's predefined reports, a sets report, for example, must contain the word 'SETS' in its name. Similarly, a pairs report must contain 'PAIRS' and an overlap report, 'MERGE'. Also, if you have found matches to business level, check you have chosen Business in the Report Format drop down list (View Matches dialog).

If you are using a report that you have modified, the report may be corrupt. It is worthwhile checking whether you have the same problem using the standard Business or Residential report, as delivered.

(c) Indexing Problems

Otherwise, it could be caused by an indexing problem. Check this by selecting Browse Imported Records from the Import menu and ordering it on UNIQUE REF. If the unique reference numbers are not progressive the Index is corrupt. Solve this by using the Database Utility option, Reindex to reindex the Main File(s).

(d) Find matches before you view them

matchIT may display completely false matches, with no similarities whatsoever if you select View Matches before you Find Matches. The reason is that matchIT retains information from the last matching runs that took place for a Main File in that directory.

Frequently Asked Questions

The Frequently Asked Questions option is located in matchIT's Help menu. This option will link the user directly to our main FAQ page at www.helpIT.com. The FAQ page is updated regularly.

Technical Support

matchIT has a 30 day warranty and telephone support included in the license fee. After that, maintenance, enhancements to the functions you have purchased and telephone support are available, at 20% of the purchase price per year.

If you are unable to find the answers you need in the Online Help or on our web site (especially in the "Frequently Asked Questions" area) or the documentation provided, please contact our support desk.

Further Information

Getting Started Guide

The matchIT Getting Started Guide is supplied to help new users become familiar with matchIT's main functions. We recommend that all new users work through this guide to gain familiarity with the matchIT environment and basic use of the system. If you don't have a hard copy, you can print the Getting Started Guide as it is available in Adobe Acrobat PDF format from the matchIT Help menu – you can obtain the Acrobat Reader software from Adobe's web site, www.adobe.com.

Online Help

More detail on some topics is given in the Online Help, to provide further information about advanced use of matchIT and the many parameters which can be manipulated. This is recommended for users who are familiar with matchIT and wish to learn about its advanced features, as well as for referencing when you want to configure matchIT for a specific task.

Additional Modules/Licenses

matchIT can be easily expanded to accommodate different requirements. Some of these additional features you may have already noticed from sections of this manual. To find the level of matchIT installed, choose about matchIT from the Help menu.

Additional user licenses are also available, at a reduced rate for each additional license.

For further information or to discuss your requirements, please contact your supplier (see "Contact Us" below).

Contact Us

US Office

Website:	www.helpIT.com
Mailing address:	helpIT systems inc 560 South Winchester Blvd., 5th Floor San Jose, CA 95128, USA
Support:	408 236 7492
Main:	866 628 2448
Fax:	408 236 7491
Support e-mail:	Support.US@helpIT.com
Sales e-mail:	Sales@helpIT.com

Appendix

Appendix A - Field Names & Usage

The section after this lists field names by their type of usage.

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
ADDRESS1		The first or only line of address	Address1 can contain the whole address e.g. with each line separated by a comma	Yes	Yes	Yes	You must have at least one address line for matching and for address validation.	Cased or replaced by data from Postal Address File
ADDRESS2-9		The second and subsequent lines of address		Optional	Optional	Optional	matchIT does not expect any structure to the address, so it is usually best to label all the address lines as Address1 through AddressN (maximum of 9 lines), even if you know that e.g. the third address line is the town and the fourth the county.	Cased or replaced by data from Postal Address File
ADDRESSEE		Free form personal name in one field, as used in addressing	MR G.C. SMITH Geoff Smith Esq. Mr G. Smith & Miss S. Brown	Optional	Optional	No	If this field mostly contains personal names, with some company names	Cased or generated from Prefix, Forenames, Surname

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		mail.					or job titles, label it Addressee. This field can be generated from Prefix, Forenames and Surname.	
ADD_KEY	8	This is a key derived by matchIT for use in matching, to pick up matches where postcode does not match or it is non-UK data. By default, it is an 8 byte field, of which the first four characters represent the phonetic key of the town/city and the second four characters the street.	"bkymlymt" for a phonetic key of "The Cottage", 8 Lime-Tree Close, Gt. Bookham, Leatherhead, Surrey. THECOTTAGE8LIMET-REECLOSEGTBOOKHA-MLEATHERHEADSURREY for a non-phonetic key.	No	Recommended	No	See "Use Fast Address Key" in the Online Help.	Yes
BASE_SCORE	10	BASE_SCORE is the lowest matching score for a matching pair or a group of matching pairs.	105 98 130	No	No	No	A group of matching records will have the same BASE_SCORE.	Generated from the matchIT Weights table
CARRIER	4	Enhances data by adding the Carrier Route Code to records run	C002 C045	No	No	No		Generated when the "Add Route Codes (CRRT)"

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		through the addressIT module.						option is turned on within the addressIT module
CASSKEY	20	A unique address key is generated for each record when it is run through addressIT. The key is generated regardless of the record being a pass or failure.	A3CGCM9FHJBFBPR1 A0ZZZZM3PPF6P564	No	Optional	No	This field is created when the "Use USPS unique address key as additional matchkey" option is turned on within the addressIT module.	Generates a unique key from the returned ZIP+4 coding which is useful in enhancing the matching keys and results
CCODE	10	Enhances data by adding County Code (FIPS) information to records run through the addressIT module.	053 125	No	No	No		Generated when the "Add County Code (FIPS)" option is turned on within the addressIT module
CNAME	50	Enhances data by adding the County name to records run through the addressIT module.	DAVIDSON LUCAS	No	No	No		Generated when the "Add County Name" option is turned on within the addressIT module
COMPANY		A company or business name.		Optional	Optional	No	matchIT can move some company names from the Addressee or an	Cased or moved from Addressee or address

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
							address line to the Company field.	
CONGDIST	2	Enhances data by adding the Congressional District to records run through the addressIT module.	05 13	No	No	No		Generated when the "Add Congressional District" option is turned on within the addressIT module
CONTACT	40 min	This is a form of the person's name that is suitable for addressing envelopes.	Mr S Brown Dr J A Baker MSc	No	No	No		Generated from the individual's name
CORRCODE	8	Provides an informational code that describes why/what was address standardized for a record.	ART3 BLRT D	No	No	No	This field is only populated if the record was successfully address standardized.	Generated by matchIT when the addressIT module is turned on
COUNTRY	30	Country	USA United Kingdom Australia	Optional	No	Optional	matchIT can move or copy countries from an address line to the Country field.	Cased or generated from address lines
COUNTY	30	UK County		Optional	No	No	matchIT can move or copy valid UK counties from an address line to the Town field.	Cased or generated from address lines

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
COY_KEY	5	Match key field consisting by default of the phonetic key of the first key word in the business name.		No	Optional	No	This field widens the choice of Match Keys if you are matching on both people and company names in the same run.	Yes
COY_NAME	40	Standardized form of the business name in upper case, used for matching.	IBM,, SMITH,KLINE,BEECHAM MINISTRY,DEFENCE,	No	Yes	No	Stored as follows: Surname, First Forename (or initial), Other forenames (or initials). Each of these will be separated by the Name Delimiter.	Yes
DEL_PRI	10	DEL_PRI is set to the Deletion Priority of each record in a match set upon deletion.		No	Optional	No	For more information, see Deletion Priorities.	Yes
DEPARTMENT		Department name		Optional	No	No	No special processing by matchIT other than casing.	Can be cased
DPBAR	3	Enhances data by adding the Delivery Point Barcode to records run through the addressIT module.	494 756	No	No	No		Generated when the "Add Delivery Point Barcodes (DPBC)" option is turned on within the addressIT

Appendix

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
								module
DPV	3	Enhances data by adding the Delivery Point Validation information generated by the addressIT module.	YN N N DN	No	No	No	Delivery Point Validation is now a requirement of CASS certification.	Generated by matchIT when the addressIT module is turned on
ERRORCODE	2	Provides an informational code that describes why a record failed the address standardization process.	8 2	No	No	No	If no ERRORCODE was generated the record is has been address standardized. The field is only populated for failed records.	Generated by matchIT when the addressIT module is turned on
FLAG	20	Populated during Import or Generate Keys for use by the Data Summary and View Records By Category menu options.		No	No	No	Each position in the FLAG field specifies a different property for that record – the properties and values are listed at the end of this section.	Yes
FOOTNOTES	8	Provides an informational code that gives the specifics of the generated DPV field.	AAN1 AAM3	No	No	No		Generated by matchIT when the addressIT module is turned on
FORENAMES		Forenames (or Christian names) or initials in one field.	J, JR, J R, John Robert, John R	Optional	Optional	No	matchIT will generate Forenames from information contained in an	Cased or generated from Addressee

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
							Addressee field if present.	
FORENAMES2		Forenames for second name.	Similar to FORENAMES	No	Optional	No	matchIT will generate Forenames2 from the second name contained in an Addressee field if present.	Generated from Addressee
IMPADD1-9		The imported address lines that were run through address standardization (addressIT).		Optional	Optional	No	If running your file through addressIT, matchIT will re-label your original address lines as IMPADD1, IMPADD2, etc..	Yes
IMPZIP		The imported Zip code that was run through address standardization (addressIT).		Optional	Optional	No	If running your file through addressIT, matchIT will re-label your original zip code as IMPZIP.	Yes
INITIALS		Initials of either just the middle forename(s), or of all the forenames.	Where the name of the person is John R Smith: JR or J R if FORENAMES is not present or empty. R if FORENAMES contains John.	Optional	Optional	No	If Forenames contains initials, they will be moved into the INITIALS field if present.	Generated from Addressee or Forenames
INITIALS2		Initials of either just the middle forename(s), or of all the forenames, for a second name.	As for INITIALS	Optional	Optional	No		Generated from Addressee

180

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
JOB_TITLE		A job title or department name.		Optional	Optional	No	matchIT can move some job titles or department names from the Addressee or an address line to the Job_Title field.	Cased or moved from Addressee or address
LOT	5	Enhances data by adding the Line of Travel Information to records run through the addressIT module.	0356A 0108A	No	No	No		Generated when the "Add Line Of Travel Information (LOT)" option is turned on within the addressIT module
MAILING_ID	10	This field will be set to the campaign identifier entered in the Output Campaign Options window, for every record included in the output file, or for which a label was printed or letter produced.		User specified	No	No	matchIT **Campaign** only	
MAIL_DATE	Date	This field will be set to the mailing date entered in the		User specified	No	No	matchIT **Campaign** only	

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		Output options window, for every record included in the output file, or for which a label was printed or letter produced.						
MATCH_REF	Same as Unique_Ref	When you Flag Matches, MATCH_REF is populated with the UNIQUE_REF of the record that is retained from every set of matching records (a "duplicate group"). Alternatively, Group Matches from the matching menu just populates MATCH_REF.	If records 124, 318 and 960 are in the same match set and following the normal deletion rules, record 318 is deemed to be the record to keep, MATCH_REF will be set to 318 for all 3 records in the set.	No	Optional	No	For records that are not part of match sets, the MATCH_REF is set to the UNIQUE_REF.

MATCH_REF is not populated by merge/purge processing unless you use the Multiple File Wizard. | Yes |
| NAME | 40 | Standardized form of the personal name in upper case, used for matching. | KELLY,JOHN,PATRICK

SMITH,G,C

TREVOR-ROPER, HUGH,

FISCHER,HANS, & BRAUN,MARINA, | No | Yes | No | Stored as follows:
 Surname, First Forename (or initial), Other forenames (or initials). Each of these will be separated by the Name Delimiter. | Yes |
| NAME1 | 5 | Standardized, usually phonetic, form | dymym

D545 | No | Yes | No | This is used as a component of Match Keys for | Yes |

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		of the surname or first key word in the company name.	DENHOLM depending on the phonetic matching option chosen.				either Personal (individual, family or contact) or Business matching.	
NAME2	5	Standardized, usually phonetic, form of the first forename, or its initial or the second key word in the company name.	Similar to NAME1	No	Yes	No	For personal matching, this can be replaced by an equivalent form of the forename e.g. Anthony instead of Tony, or a phonetic key of Anthony.	Yes
NAME2-FOUND	1	An indicator of whether the first forename in the Input record is in the NAMES table.		No	Yes	No	If the first forename in the Input record is in the NAMES table for the sex, NAME2FOUND will be set to "Y". If it seems to be inconsistent with the SEX or PREFIX for that record, it will be set to "X".	Yes
NAME3	1	Standardized, usually phonetic, form of the second forename, or its initial or the third key word in the company name.		No	Yes	No	Defaults to one character wide so it matches "middle names" just on their initial.	
NAME_KEY	6	Match key		No	Optional	No	This field makes	Yes

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		field consisting by default of NAME1 plus the first character of NAME2					the choice of Match Keys simpler.	
POSTCODE	8 min	Usually contains a full UK postcode.	GU14 7BQ W1A 1AA E3 5HD KT228DY	Optional	Optional	Yes	matchIT can extract (i.e. move) valid UK postcodes from an address line to the Postcode field.	From address lines or from Postal Address File
PREFIX	10 min	Personal title	Mr Mrs Dr Professor	Optional	Optional	No	matchIT will generate Prefix from information contained in an Addressee field if present; otherwise it will try to deduce it from the forenames if present i.e. Mr or Ms.	Cased or generated from Addressee
PREFIX2	10 min	Personal title for second name	Similar to PREFIX	No	Optional	No	matchIT will generate Prefix2 from the second name contained in an Addressee field if present.	Generated from Addressee
PREMISE	10 min	The premise number part of the address – this is the building number, not the apartment or unit number	260 2A 12-14	Optional	Optional	No	matchIT can move or copy some premise numbers from an address line to the Premise field.	Cased or generated from address lines

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		if there is one.						
QUAL	10 min	A qualification or qualifications.	B.Sc., ARCS, BCom	Optional	No	No	matchIT will generate Qualification from information contained in an Addressee or Surname field if present.	Optional
QUAL_ADDR	2	Address quality score	3 9 5	Optional	No	No	matchIT assigned quality score for address information. See Address Quality on page 144.	Generated from address information.
QUAL_ORG	2	Organization quality score	99 50 05	Optional	No	No	matchIT assigned quality score for Organization information.	Generated from Organization information.
QUAL_INDIV	2	Name quality score	8 1 6	Optional	No	No	matchIT assigned quality score for Name information. See Name Quality on page 145.	Generated from Name information.
QUAL_EMAIL	2	Email quality score	9 0 2	Optional	No	No	matchIT assigned quality score for Email information. See Email Quality on page 146.	Generated from Email information.
REST_PHONE	10 min	This is the local part of the telephone number, obtained by	360070 in (01372) 360070	Optional	Optional	No		Can be derived from the TELE-PHONE field. It is useful as

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		stripping out the area code (Std_code), spaces, punctuation characters and characters after the local number such as extension numbers.						a match key either on its own or in conjunction with other fields
SALUTATION	40 min	This is a form of the person's name that is suitable for the top of the letter.	Dear Mr Brown	No	No	No		Generated from the individual's name
SET_DUPS	5 numeric	If flagging match sets, SET_DUPS will contain the number of duplicates in a match set i.e. the number of records in the set less 1.		No	Optional	No	This field will be set to 0 if the record is unique, > 0 if the record is part of a match set.	Yes
SEX	1	The gender of the personal name, if present.	M, F or blank	Optional	Yes	No	matchIT uses this field in Individual matching if the corresponding Matching Option is set.	Generated from PREFIX or FORENAME, if not provided
STATE	30	Non-UK State or Province.		Optional	No	No	matchIT can move or copy non-UK states that are in its look-up table from an address	Cased or generated from address lines

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
							line to the State field.	
STD_CODE	7	This is the area code for the telephone number.	01372 in (01372) 360070	Optional	Optional	No		Can be derived from the TELEPHONE field
SUFFIX	10 min	A title in a separate field following the surname.	Esq, Jr	Optional	Optional	No	matchIT will generate Suffix from information contained in an Addressee or Surname field if present.	Optional
SURNAME		The surname, typically when the name is split into component parts of Prefix, Forenames and Surname.	Smith, van der Valk	Optional	Optional	No	matchIT will generate Surname from information contained in an Addressee field if present.	Cased or generated from Addressee
TELEPHONE		Telephone number	01372 360070 (01372) 360 070 01372-360070	Optional	Optional	No	matchIT can split telephone numbers into two parts: STD_CODE for the area code and REST_PHONE for the local part of the number. The user can then effectively use REST_PHONE for matching.	No
UNIQUE_REF	15 max	This field must		Optional	Yes	No	If you don't have	Optional

Field Name	Width	Description	Examples	Input	Matching	Mailsort	Notes	Generated
		be a unique key to the imported Main File record and must be of character type.					unique reference numbers in the input file, matchIT allocates them sequentially.	
ZIP	10 min	Usually contains a non-UK zip or other postal code.	US: 60010, 60010-5238 Australia: 2060 Canada: N1P4N4	Optional	Optional	No	matchIT can move or copy some non-UK zip (postal) codes from an address line to the Zip field	From address lines

Appendix B - Flag Field

Pos	Flag Name	Values
1	Exclusion Status (see Note at bottom of the table)	**X** if exclusion words are found, **blank** otherwise
2	Company Extracted	**C** if company name has been extracted
3	Company Acronym	This is set to **1,2** or **3** if any part of the company name is an acronym
4	Job Title/ Department	**J** if job title extracted, **D** if department extracted
5	Verified Postcode	**V** if postcode was verified OK
6	Extracted Postcode	**E** if postcode was extracted OK
7	Reserved	
8	Generated Prefix status	**P** - generated prefix, **Q** - changed prefix, **S** - used supplied prefix **blank** - no personal name processing
9	Salutation status	**S** - default salutation generated **G** - non-default (success!) salutation generated **blank** - no salutation field
10	Original Initial	Set to the first letter of the input forename field
11	Original Initial #2	First letter of second input forename, if any
12	Second Name Sex	Set to sex of second name if supplied & possible to calculate
13	Foreign status	**F** if record is foreign
14	Postcode vs. Town Flag	**X** if postcode disagrees with post town
15	Key Generation Method	If Phonetic Address Key was used: **B** - business keys used, **I** - personal keys used If Fast Address Key was used: **C** - business keys used, **J** - personal keys used
16	Premise status	**X** if premise extracted,

Pos	Flag Name	Values
		C if premise copied
17	Zip status	**X** if zip extracted, **C** if zip copied
18	Town status	**X** if town extracted, **C** if town copied
19	County status	**X** if county extracted, **C** if county copied
20	Country status	**X** if country extracted, **C** if country copied.

There is a setting in Basic Parameters dictating how to treat exclusions. If it is set to Delete, these records will be deleted on Import. If it is set to Mail, they will not be deleted, but will have an 'X' placed in their Flag field.