Importing Data into TMG

by Ben Franklin

Perhaps most TMG users come to TMG from another genealogy program. That is to say, you have already entered information into your previous genealogy program, so you want to import your genealogy data from that software. In addition, if you are collaborating with fellow researchers, you may want to import data created in many different programs on an on-going basis.

TMG has two strategies for importing data: One approach uses Genbridge™ technology to import data directly from most current genealogy programs, so this can be the best way to quickly and accurately import data. The other approach is to import using GEDCOM data.

The purpose of this presentation is to help you get the best possible result of an import and provide guidance identify and clean up remaining issues.

Basic Steps

1. Begin with the Simple Wizard. This first iteration will be a benchmark or experiment. You might want to write down the name of the data set, and what options you used. For this first import, you should use the default values.
2. Inspect the data. Don't forget to look at the sources and various event types.
3. Return to Step 1 and use the Advanced Wizard. Change the defaults, based on the results of Step 2. You might even want to change the information in the source program, if you can. Repeat until the import is as accurate as you can make it.
4. Filter out unwanted individuals, if needed.
5. Flag all of this data as "NeedsCleaning" with a default value of "Y"
6. Manually inspect and fix every person in the database, marking off the Flag created in Step 5.
7. Merge the entire data set into the target data set, if desired.
8. Merge duplicate people.

Some Preliminary Information
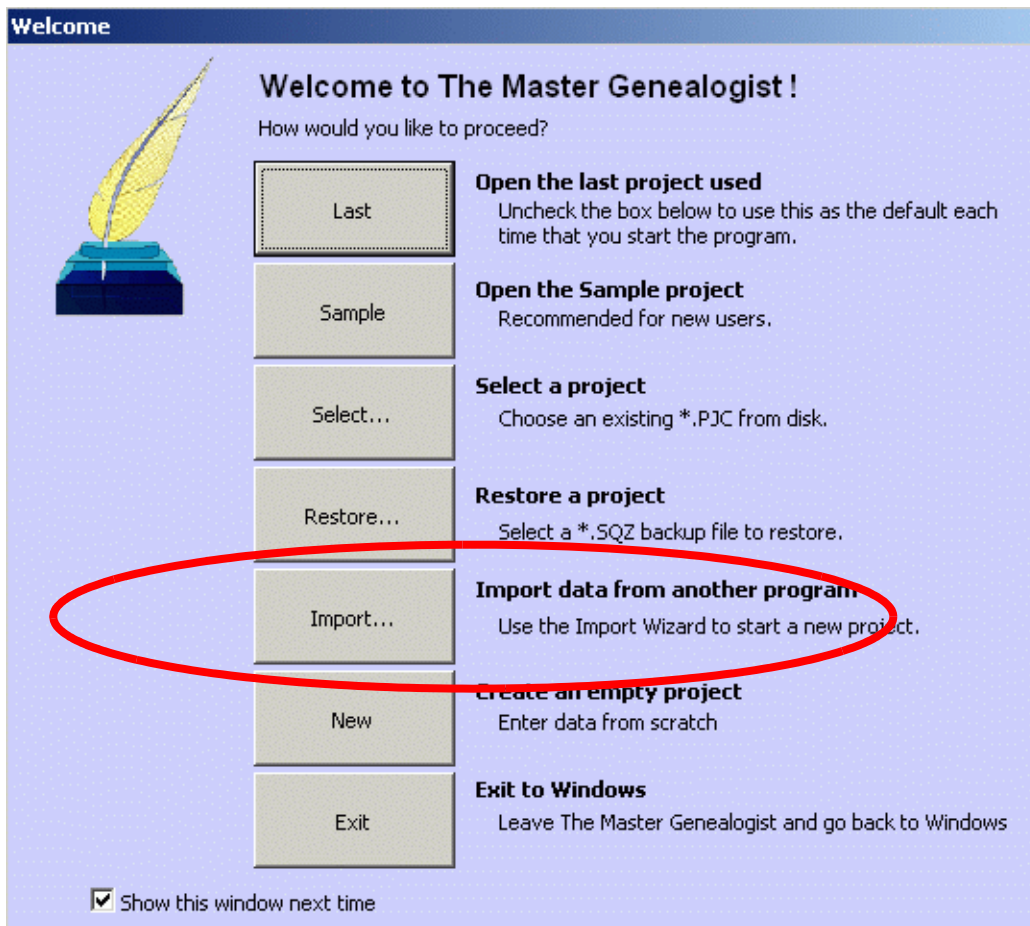
Prerequisite knowledge

In order to understand this presentation and place the information in context, you will need to know a few TMG concepts. Many of these topics have been dealt with in detail in previous presentations. These include:

- Master Place List
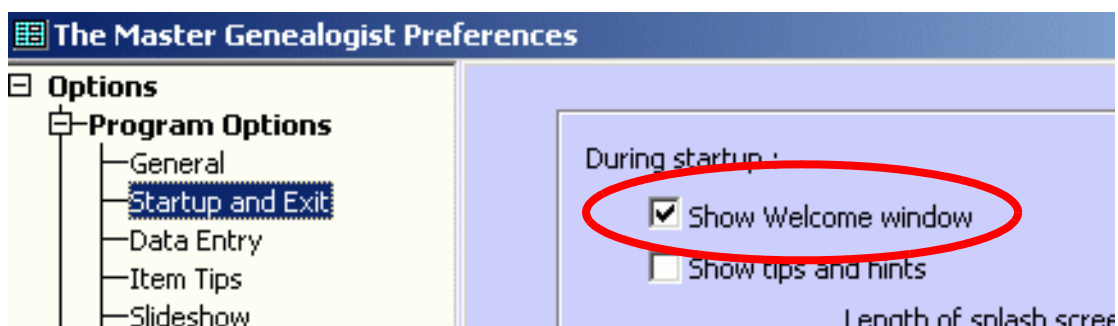- Sources
- Flags
- *TMG Utility* by John Cardinal

You will need to be familiar with all of the above in order to determine if you want to change data in the original program or in TMG. You will probably need all of these tools at some point, and you will need to know how to use TMG in order to know what defects are easy to fix after import.
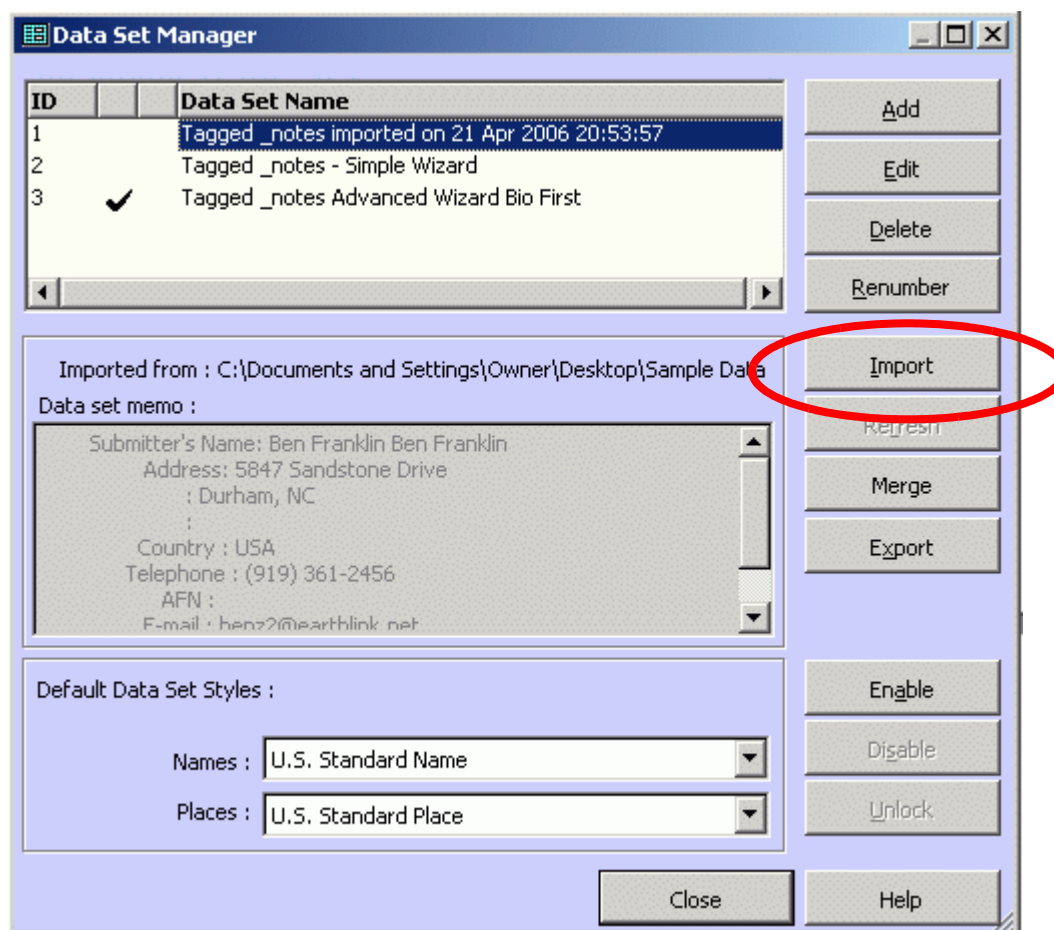
Known Import Bug

CAVEAT: There is a known bug that may occur when you are using **File > Import**. Until this bug is fixed, the way to get around this is to exit TMG and select the import option on startup, or start in the Data Set Manager, as described below.

Welcome

Welcome to The Master Genealogist !

How would you like to proceed?

Last — Open the last project used
Uncheck the box below to use this as the default each time that you start the program.

Sample — Open the Sample project
Recommended for new users.

Select... — Select a project
Choose an existing *.PJC from disk.

Restore... — Restore a project
Select a *.SQZ backup file to restore.

Import... — Import data from another program
Use the Import Wizard to start a new project.

New — Create an empty project
Enter data from scratch

Exit — Exit to Windows
Leave The Master Genealogist and go back to Windows

☑ Show this window next time

In order to see the above menu, you must have the "Show Welcome Window" option selected in Preferences. This is easily "unselected" from the Welcome Window itself, so you may have this disabled. You will need to enable it.

The Master Genealogist Preferences

□ Options
 □ Program Options
  — General
  — Startup and Exit
  — Data Entry
  — Item Tips
  — Slideshow

During startup :
☑ Show Welcome window
☐ Show tips and hints
Length of splash screen

To specify an existing project, and thus make the data available as another data set in that project, you must import the data using the Data Set Manager.

Lastly, if the above-mentioned bug doesn't stop you, you may use the **File > Import** menu to import into the current project, or create a new project.

What is GenBridge?

Before we start talking about importing in TMG, we should probably discuss the term "GenBridge." Some people are confused by the term, thinking that it is a separate program, or a special menu selection in TMG. Genbridge is simply the part of TMG that does the reading and importing data from other genealogy programs in their native formats.
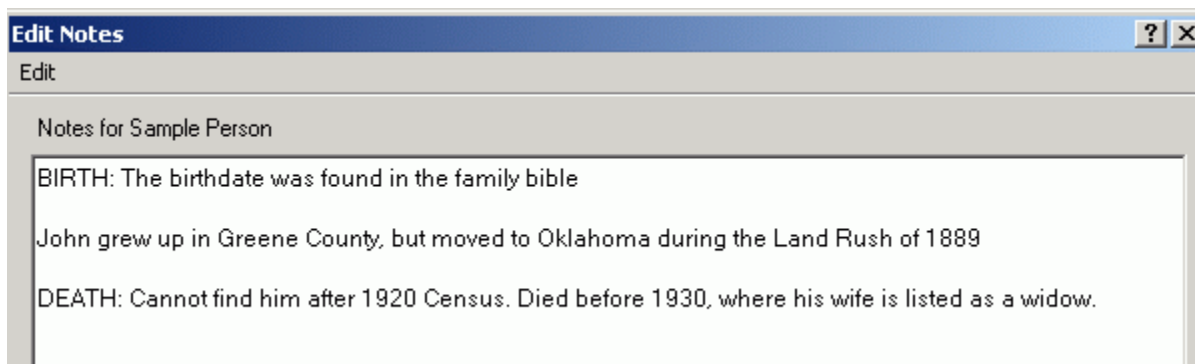
Why Doesn't TMG Import "Accurately"

The bad news is that there are several reasons that an import of data (using default settings) will most likely **not** give you the results that you wanted or expected. First, TMG's data model the most complex and detailed of any genealogy program available. Sometimes these complexities do not map well into the more primitive, simplistic data models of other programs. Second, many researchers have their own unique, unconventional methods of recording data in their various programs. Last, but not least, if you are using GEDCOM, remember that there is considerable variability in the way genealogy software has interpreted the standard that can cause the data to end up in TMG in an unexpected way, despite the fact that the other program that output the data did so in a way that is fully compliant with the GEDCOM standard. The original purpose of GEDCOM was to convey established genealogical facts. Because it was not designed to

facilitate on-going research projects which might contain ambiguity, incomplete data, etc., it is not the best way to store, transfer or share any genealogical information other than the most basic facts.

The good news is that TMG has many options to improve the accuracy of your import. More good news is that you can use *TMG Utility* to clean up some types of data import problems.
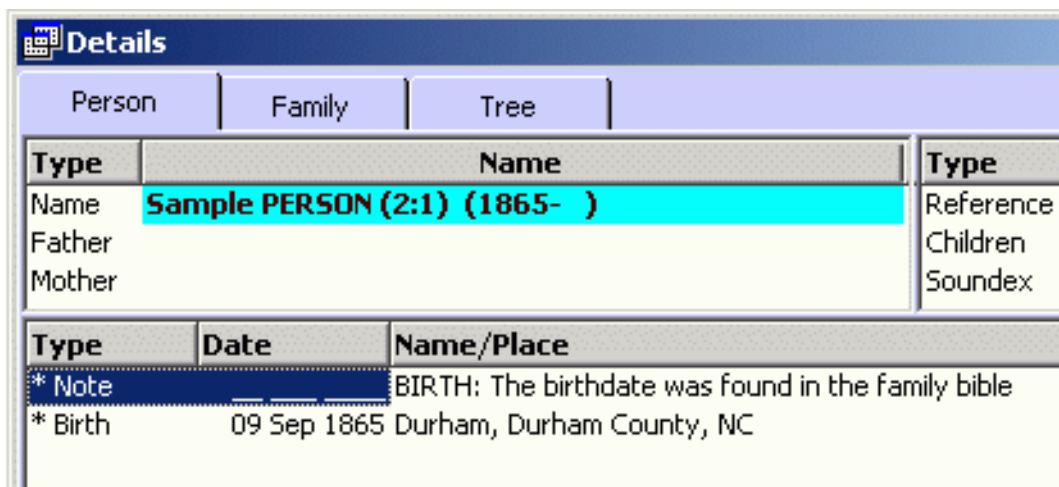
Interpreting Tagged Notes

Many programs do not associate notes with a specific event, but rather have just one note that applies to the individual as a whole. During import, GenBridge or GEDCOM import has the ability to interpret portions of these notes as tagged notes, if you select the "Interpret Tagged Notes" option. That means that if you have a note associated with the individual that looks like:
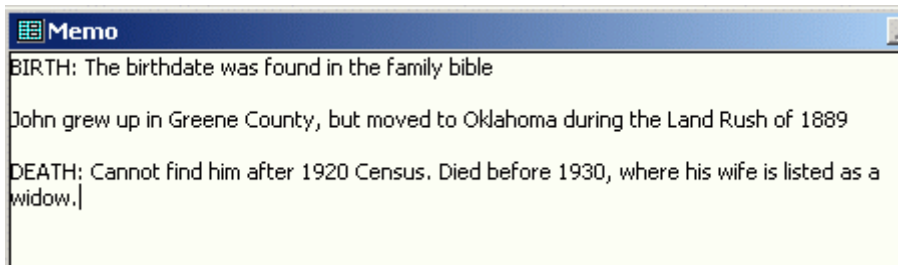
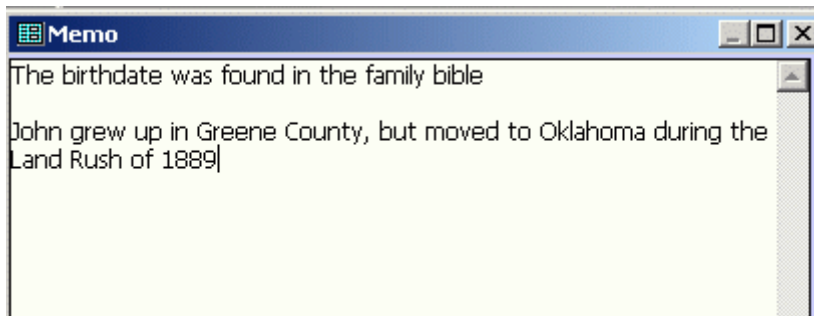

Using the Simple Wizard, in this case, produces two tags:

:

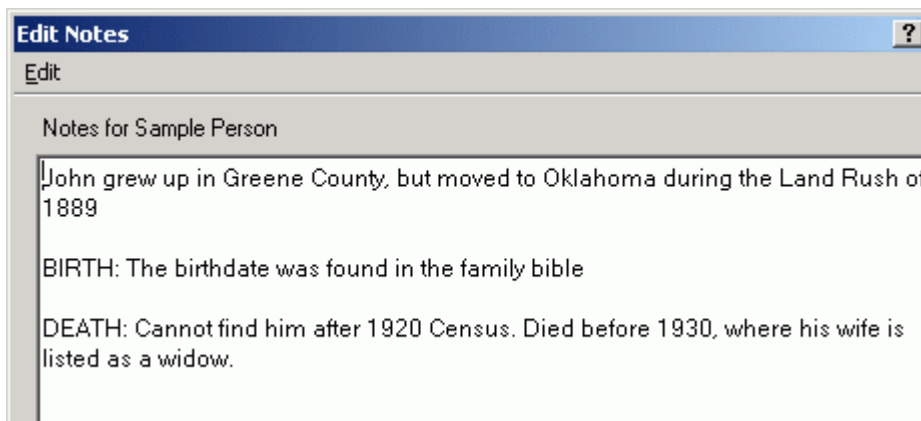...and places the entire note in one Note Tag:

Using the Advanced Wizard, when you select "Interpret Tagged Notes", there will only be two Event tags, Birth and Death. The Birth note will include the biographical information, thus:

Notice that TMG interprets all of the individual note data as tagged, and places the biographical information ("John grew up...") in the Memo for the Birth Event Tag. Note that the tag ("BIRTH:) is not included in the Memo.

In PAF, moving the biographical information to the top of the note:

...Using the Advanced Wizard with "Interpret Tagged Notes" selected:

Details

| Person | Family | Tree |

Type	Name		Type	Valu
Name	**Sample PERSON (3:1) (1865-)**		Reference	
Father			Children	0
Mother			Soundex	P625

Type	Date	Name/Place	Ag
* Death	_ ___ ____	Cannot find him after 1920 Census. Died before 1930, where his	
* Note	_ ___ ____	John grew up in Greene County, but moved to Oklahoma during the	
* Birth	09 Sep 1865	Durham, Durham County, NC, The birthdate was found in the family	

...notice that in this case, there are now three Event tags and the biographical information that was not tagged is placed in the Note tag.

Because of this, you might prefer to go back to your old genealogy program and add tags to the Notes, or otherwise edit them. Alternatively, you may prefer to **not** select the "Interpret Tagged Notes" option if you plan to re-edit all of the notes anyway, and thus all of the notes would be in the same Event Tag and would appear as originally entered.

FTM's "More About" field, or PAF's Individual Notes: Unless they are tagged notes (as described above), notes are imported into TMG as a single Note Event tag for each person, with the contents placed in the Memo field. Depending on just how the data was originally entered, it will probably include a variety of information, including biographical data, source data or even research notes and suggestions. Diligent users of FTM or PAF will also include source references for the person, and sources for the facts included in the actual text of the note, because sources cannot be attached to this field in those programs.

A Note which refers to an entire individual may include several separate topics. In TMG it may be best to break these into two or more separate tags, with their associated dates and places. Not only does this allow you to see them in the Details window in the correct chronological sequence, but it allows you to control which tags are printed in reports if you use different tag types for different types of data. One way to do this is to make multiple copies of the tag (use the **Add > Copy Tag** menu) and then edit each copy, deleting all but the information on the current subject. You can change the Tag Type and transfer the date and place information to those fields if desired.[1] If there are source notes embedded in the text, you may want to replace them with TMG source citations.

This will result in several issues in TMG. First, such notes are generally written as narrative sentences. However, the default behavior of TMG's Note tag the sentence structure adds the name of the subject, and any date or place information when creating narrative reports. The result is something like: "He John was raised in a small town..." The pronoun "He" would have been added by the Note tag's sentence structure. There are several possible solutions. The best may be to change from a Note tag to another, more appropriate Event type, and edit the Memo to remove the person's name from the text, and copy any embedded date

---

1. In order to help preserve data integrity, TMG restricts the ability to change tags to within the same Tag Group. When you need to change an event to a completely different type, you can use *TMG Utility* to change the tag type.

and place information to the appropriate fields. But if you are otherwise satisfied with the flow of the note, you could just change the Tag Type to Anecdote, which simply prints the text of the Memo field, and add a sort date to position the tag correctly in the events list of the Person View, Individual Narrative report, etc.

Keep Track Of Your Progress With Flags

If you are going to edit or cleanup the data over an extended period of time, you will need some way to keep track of what you have finished and what still needs to be done. In TMG the easiest way to help keep track of this is to create a custom Flag. In fact, you may want to define several flags for this purpose if you work on several types of cleanup or validation projects at the same time. Alternatively, you might create a flag for each item on your cleanup checklist, and clean your data set item-by-item, rather than person-by-person. When you are done with a flag, just delete it.

When you create a flag, you can assign it a range of possible settings, the first of which is the default. Thus, if you were to create an "NeedsCleaning" flag, you would create it to have a default value of "Y" and an alternate value of "N". Then as you edit each person in the data set, you would change the "NeedsClean-ing" flag for that person to "N." As you progress through the data set, you can use the Custom Report Writer to generate a "List of People" report based on a value of the "NeedsCleaning" flag, you can use a Filter in the Project Explorer to focus on the people that need to be edited, and as you display the Person View of an individual, you can see if you have cleaned up this person.

To create a flag:

* Open the Flag Manager (use **File > Flag Manager**)
* Click the Add button.
* Name your new Flag and assign it possible values. The first value will be initially assigned to every individual in the data set when you close the Flag Manager.
* Close the Flag Manager.
* Reopen Flag Manager, select your new flag, and click Edit.
* Reverse the original order of the values to change the default. This will not change the values set for the existing people, but will change the default for any new people you add, who will not need to be "cleaned" because you are adding them manually.

In order to display or edit the Flag values for each person, you will need to view the Flag Window. It is not part of the default layout that comes with TMG. Display it by clicking the Flags Window toolbar icon – – or use the **Windows > Flags** menu. If you want it to always be open, adjust the size of the other TMG windows to make a room for it, and save your new layout.

To change the value of your "NeedsCleaning" Flag for a individual you are finished with, first double click on the Flag's name in the Flag Window. When the Edit Flag window appears, double click on the new value in the left column.

Accent Color

You may also want to set an Accent color so you can quickly spot any individual who needs cleanup. Many TMG users employ the Accent colors to keep track of various characteristics. You may want to make the "NeedsCleaning" Accent the first choice, using a unique set of colors, so that until cleanup is done, normal Accent colors are overridden.

First, Try Simple Wizard

Import the data using the Simple Wizard. At this point it would probably be best to think of this as an experiment to see what options you might want to change in the Advanced Wizard. Use the native data for-mat if available. The resulting data set will usually be more accurate than GEDCOM import, resulting in less data loss and cleaner data after import. The following programs are supported:

Family Gathering™, Family Origins™ (v4 - 10), Family Tree Maker™ (v3-11, 2005), Generations™, Legacy™ (v3-5), Personal Ancestral File™ (v2 - 5), Roots IV™, Roots V™, Ultimate Family Tree™, Visual Roots™.

Remember that if your program is not listed here, first contact Wholly Genes support <support@wholly-genes.com> to determine if they can help. They may be able to perform the import for you and achieve a better result than can be achieved via GEDCOM. For more information about GenBridge, see:

`http://www.whollygenes.com/tmggenbridge.htm`

You can create a test data set where you can try out various import options to find those that work best for you, without worrying about interfering with your real research data set. You may want to use your other genealogy program to create a simple database with sample data to import, or you prefer to just use your real data to explore alternate methods with the data you will actually use. Just be sure to give your new TMG data set a title that indicates to you what it is, so you don't accidently mistake it for a legitimate research data set.

Start by importing using the Simple Wizard and all defaults. Write down the name of the data set you have created and what options you have selected.

Inspect The Data

The TMG Import Advanced Wizard includes a number of options when importing. Because of differences between how TMG and your previous program records your data, and because of the variance in methods used by the users of each program, you will probably obtain a more accurate import by changing the options. So, before investing time in editing of your imported data, examine it closely. This step is frequently overlooked, skipped or shortened. Most people want to jump into changing their new TMG data set. **Don't !** Take a deep breath and step back from the keyboard! You **must** inspect your data. Remember that it is much, much easier to adjust the import options and try again than it is to discover later after you have a lot of changes to your TMG data, you wished you'd imported it differently. It is very likely that when you inspect it, you will find that you should have chosen a different set of options. If you go back and re-import, any changes you have made will be lost. You must be extra sure that you have already achieved the optimal import before you start modifying your data.

Create some sample reports and study them. If the results are less than ideal, consider making another import with different settings. A little experimentation with different settings may save you many hours of cleanup later.

Remember to not only inspect a number of individuals, but also look to see how the sources were transferred. Look at the place names in the Master Place Name List. Place names might be easier to clean up in TMG, using the edit function in the Master Place Name List rather than in the original program, or by changing Import options.

Be extra careful to see that TMG has transferred the various event tags correctly. For some tags, it is simple to change them to another tag within TMG. Other tags may require the use of *TMG Utility*. You might want to print out a "List of Events" report.

Also look at the information contained in the "Cleaning Up Imported Data" portion of this presentation. It lists a number of issues that you can expect to find.

You might want to write down the defects that you find so that you can remember what needs to be changed before re-importing.

Re-Import with Different Options

Hmmm... This might be the point at which you wish you had written down the name of the data set and the options you used last time... Using the Advanced Wizard, select different options to correct the problems that you found when you inspected the data.

Only a few options are available with the Simple Wizard, but all options are available only if you choose the Advanced Wizard at Step 2 in the Import process. When importing from FTM in particular, you may obtain better results with the Advanced Wizard, especially if the FTM user has made extensive use of FTM's "Facts" feature.

FTM Facts consist of a date field, and a single combined field called "Comment/Location" which contains a mixture of notes and places. Depending on the type of fact involved, users can enter either a comment or location information, or even both. The Import Wizard will attempt to place the contents of the FTM Comments/Location field in the correct field in TMG, but may not always be successful. This is especially true when the user has been inconsistent, putting location information in this field in some cases, and comments in other cases. You might be able to improve the accuracy of the import by making choices in Step 5 of the Advanced Wizard. (Note: TMG **cannot** separate Comments and Locations in the same field when they are delimited with the slash character, as recommended in Help for more recent versions of FTM.)

Maintaining A Multi-import From The Same Author Of Updated Materials

You may want to collaborate with another user who is currently also updating their research database. If this is the case, there are several ideas that may help you:

- Create a Flag that identifies this data. You can then set an Accent color for these people, filter them out, etc.
- Keep the data in a separate data set in your project, so that you can overwrite the entire data set, turn it off, etc.
- This may be the only circumstance where you will may want to import this as a **Locked** Data Set. Leaving it locked will remind you if you forget and try to change the information found in it.

Cleaning Up Imported Data

There are several approaches you should carefully consider before you clean up your data. Which is best for you depends on how many individuals are in your data set, how anxious you are to get everything tidied up, and how fully you plan to apply the advanced features of TMG. If you want to exploit the advanced source recording features in TMG, you will need to examine the alternatives and decide which method best serves your purposes before spending time cleaning up your sources. Likewise if you use the Witness and Role features to improve how you record census records, wills, and similar documents that list multiple participants, you may want to decide just how you want to use these features before you start to clean up each person in your data set. Alternatively, you may choose to get everything tidied up using TMG's basic features before applying the more advanced functions.

Sources

This presentation is not specifically about sources, so I will deal with them rather briefly. In my experience, TMG offers more options for recording source information than any other genealogy program. As a result, while all the information recorded in your previous program might be successfully imported in some form into TMG, it is unlikely to be in a format to take advantage of all of TMG's features. Depending on how the data was entered in the previous program, it may need extensive editing just to print legible reports.

You may prefer to edit all of the source information immediately after import, to show in reports as desired. Alternatively, you may want to defer this task until you have explored the various ways to enter sources and have found the approach that best serves your needs. You might decide that the move to TMG is a good time to re-evaluate your overall approach to sources, re-verifying each source entry. Select the approach that best suits your needs.

Master Place List

As with other aspects of the TMG data model, place names in TMG are more detailed and flexible than the paradigm found in other programs. Typically, you enter address, city, county, and state each in separate elements of the place in TMG. This allows you the flexibility to configure reports to display the entire place name, or any specific parts of it in various reports.

The Import Wizard interprets the various parts of the imported place names to assign them to the correct field in TMG. In many cases, particularly those that contain the US states (or their abbreviations), it does a pretty good job, but other cases will confuse it. Any defects in entering place names, such as: inconsistencies, incorrect spelling, etc., can be addressed in the Master Place List. The Master Place List can be a powerful tool for correctly aligning the place elements and standardizing places names, so that you don't have to edit the places in each Event tag individually. Changes entered in the Master Place List are made to every Event tag which uses that exact place name.

If you find non-place data in the Master Place List (for example because of problems with combined Comments/Location data in FTM) select that element and click the Events button. This will display the Master Event List, from which you can access and edit the specific tag that contains the misplaced information. (Note that after you make the correction, the original entry will still be displayed in the Master Source List until you optimize your data set, using **File > Maintenance > Optimize**)

Create An Inspection Checklist

You will need to review every person in your data set in order to take advantage of TMG's strengths. Fortunately, there is no need to accomplish this all at once. So the question arises, precisely what should you look for as you edit each person? Like any other research project, you will need to approach this "with the end in mind." Thus, you might want to inspect the data using the various reports that you normally use, or using reports that you plan on using in the future. I find the Individual Narrative useful for this. The final answer will depend on how you plan to use various TMG features, and your own data entry methods and preferences.

The list below contains some suggestions, but you may have a different list, depending on your preferences, the program from which the data was imported, and the methods used to enter data in the source program.

- Insure the name is entered according to your standards (unknown names left blank, nicknames in a separate Name tag, etc.)
- Verify that the flags are set correctly. Set the Living Flag and any other custom Flags that need to be set. You may want to do this with *TMG Utility* or using a filtered list.
- Add source citations for parent relationship tags.
- Verify that sources are cited for all event tags.
- Add sort dates to undated events so that they are correctly sorted (See Note 1).
- Add a Married Name tag for married women so that they also appear by married name in the Picklist The married name will be generated automatically when you edit the Marriage tag (if you have that option set in Preferences), but when you Import, it doesn't get created.
- If multiple items will appear in a single tag, separate these into appropriate discrete Event tags. Refer to the "Interpreting Tagged Notes" section of this presentation.
- Change any source notes embedded as Memos in the various event tags Memos into correct TMG-style citations.
- Examine "double date" entries and confirm that the date was correctly entered, if possible.
- Assign sureties to every source, if you use these. (See Note 1).

Note 1: Requires "Advanced Data Entry Mode". This is enabled in **Preferences > Program Options > Data Entry**

PAF does not force all dates into a standard format, although it warns you if you enter something non-standard. FTM does force dates into a standard format. Other programs may not, and may allow a wide variety of date formats. TMG will recognize many different date formats during import and convert them to one of its standard formats. The date formats that it doesn't recognize will be converted to text and accepted as "irregular" dates. They are not interpreted as dates by TMG. If there are such dates in your imported data, this should be included in your checklist for review to see that they are treated as you prefer. If nothing else, you will want to enter regular dates in the sort date fields so that the tags will appear in chronological order.

Double Dating

No, this has nothing to do with socializing(!) There is a particularly bad practice in WFT (World Family Tree) submissions which were sent in with blank dates. The WFT process automatically adds estimated dates (WFT Est. xxxx-xxxx). TMG Import interprets these dates as legitimate double dates. ("between xxxx and xxxx"). All of these should be deleted.

Additionally, FTM automatically accommodates the crossover between the Julian and Gregorian calendars. If the complete double date is entered in FTM, it should be correctly stored in FTM. However, if it was entered as only a single year, FTM will replace that with double date, interpreting the year entered as the second year. If that was not correct, the date in FTM is now incorrect. TMG will accept either single or double dates for this period, and will preserve them as they appear. If you have such dates in your imported data and don't know if they have been correctly entered, while you are cleaning up that data, it may be a good time to correct these to agree with your research, if possible.

Merging People

Properly merging people in TMG is a very, very laborious and tedious process. Merging the individuals themselves is pretty quick, but merging the event tags (and their source citations) is all completely manual. If you want to maintain source references to each of the tags, you will have to combine each one manually.

In fact, it is sufficiently time-consuming that you might prefer to do merging and other data manipulation in an application that does it more efficiently (which may be **all** other genealogy programs). This is one of the primary reasons that I continue to use PAF in conjunction with TMG. I use PAF to manipulate the information, then I import it into TMG.

Fix Problems in the Original Program

Some problems will be easier (**much** easier) to fix in the original program, if that is available to you. This is particularly true when it is a program that you are familiar with. For example, PAF allows you to add a parent or connect children in a few seconds, whereas TMG takes minutes. You may want to do that in PAF, before import.

FTM allows siblings to be entered without entering either of their parents. TMG requires at least one parent be present to connect siblings. If your database contains such cases, you may find it much faster and easier to add a parent in FTM before performing an import rather than afterwards.

In some cases, the tools in TMG may make it easier to deal with defects after import. As mentioned previously, TMG's Master Place List generally makes the task of alignment of place elements easier to accomplish after import. Trial and error is probably the best approach to arrive at a process for specific cases.

Merging Data Sets

For the reasons that I have pointed out in previous meetings, you **cannot** merge people across projects.

Note that if you are collaborating with someone else who is actively doing research, and you plan on importing multiple times as that person send you updates, you will want to make sure that all of the information that is sent to you can be completely wiped-out each time you import. What I suggest is that you keep the information in a separate data set, and then completely overwrite it each time. If you merge this

data into your own data set, the data will be more difficult to identify and eliminate (or overwrite). If you start merging individuals, the data will be, for all intents and purposes, inextricable.

Legacy Import Note

The answer is to try both direct import and GEDCom and compare the results. There are a few omissions from the direct Legacy import such as the citation and source 'actual text' fields. The loose ends for the direct import have been reported.

Suggested Reading

The GEDCOM Standard, Release 5.5
http://homepages.rootsweb.com/~pmcbride/gedcom/55gctoc.htm

Lee Hoffman's TMG Start-up Tips (ver 5)
http://www.tmgtips.com/startup.htm

Terry's TMG Tips: Importing Your Data from Another Genealogy Program
http://tmg.reigelridge.com/Importing.htm

Terry's TMG Tips: Merging Duplicate People
http://tmg.reigelridge.com/merging-people.htm

Case Study: Transferring data from Ultimate Family Tree™ (UFT) with GEDCOM and GenBridge™
http://www.whollygenes.com/files/uft2tmg.pdf

TMG Utility by John Cardinal
http://www.johncardinal.com/tmgutil/