

Improving Your Use of FamilySearch: Data Cleanup Strategies

Geoffrey D. Rasmussen

Geoff@LegacyFamilyTree.com

Researchers should be careful when publishing/sharing information from their family file with FamilySearch, or with any online database. If the information in your family file has errors, duplicate information, or inconsistent data, when that information is published/shared online, the online version will contain those same errors, duplicate information, and inconsistent data. A casual commitment to your personal data entry standards often results in negative unforeseen consequences, especially when publishing your data online.

For this reason, before interfacing with FamilySearch's new Family Tree system, we should all take a closer look at our data.

Consistency in your locations

Take a look at your Master Location List in Legacy (View > Master Lists > Location). Does it contain more than one variation of the spelling of a place name? For example, in Legacy's sample family file, we intentionally spelled Minneapolis with five different variations to help you see a potential problem:

- Minn., Henn., MN, USA
- Minneapolis, Henn., MN, USA
- Minneapolis, Hennepin County, MN, USA
- Minneapolis, Hennepin, Minnesota
- Minneapolis, Hennepin, Minnesota, USA

Publishing the inconsistent locations may eventually result in you/others not successfully finding what you are looking for because what you are searching for, and what is in the database are two different variations of the same place.

FamilySearch's new Standard Finder helps you achieve standardization in the way you spell your locations. Its goal is to help achieve consistency in how everyone records their locations. It is not a perfect system yet. For example, we know that correct data entry standards suggest to record the place as it existed at the time of the event. Standard Finder does not yet do a great job with correctly interpreting historical locations. But it has a great start.

Assignment #1 - standardize your locations

Clean up the duplicates and variations of your place names in your Master Location List. Legacy makes this simple to do. We've created a video for you to see the step-by-step instructions on how to do this. Watch the video here: <http://tinyurl.com/yd7c8wd>

Assignment #2 - clean up your country/county names

If these locations are not immediately obvious to you:

- Victoria, BC
- Carrick, LDY
- Aberdeen, SCT
- Paris, Bear Lake, ID

then the researcher has not followed the golden rule of data entry:

Enter your locations so that there will be no misunderstanding by other researchers as to the location you are trying to express.

Legacy Family Tree's Expand/Contract Location Parts tool makes it easy to correct common abbreviations, and for the United States, to add the country name to the end of its locations where it is missing. This tool is found by navigating to:

View > Master Lists > Location > Options > Expand/Contract Location Parts.

In the Parts to Work On section, place checkmarks next to the desired countries to "unabbreviate" the various country and county names.

For United States researchers, use the options in the USA Country Name section. To prepare for full compatibility with FamilySearch, select the Add ", United States" option. This will add this country name to the end of all U.S. locations. To clarify, it will not add "United States" to the end of Bristol, Gloucester, England. If you previously typed "USA" instead of "United States" this option will change "USA" to "United States". Now click the Continue button, and Legacy will update your list.

Potential Problems tool.

Before you publish your data anywhere, use Legacy's Potential Problems tool. Using this tool, Legacy runs through your family file and checks the information for some common and sometimes unnoticed discrepancies.

In Legacy, the Potential Problems tool is located by going to Tools > Potential Problems.

The first time you use the Potential Problems tool, you should have it check all the records in the family file. To do this, make sure that on the Records tab, the first option is selected:

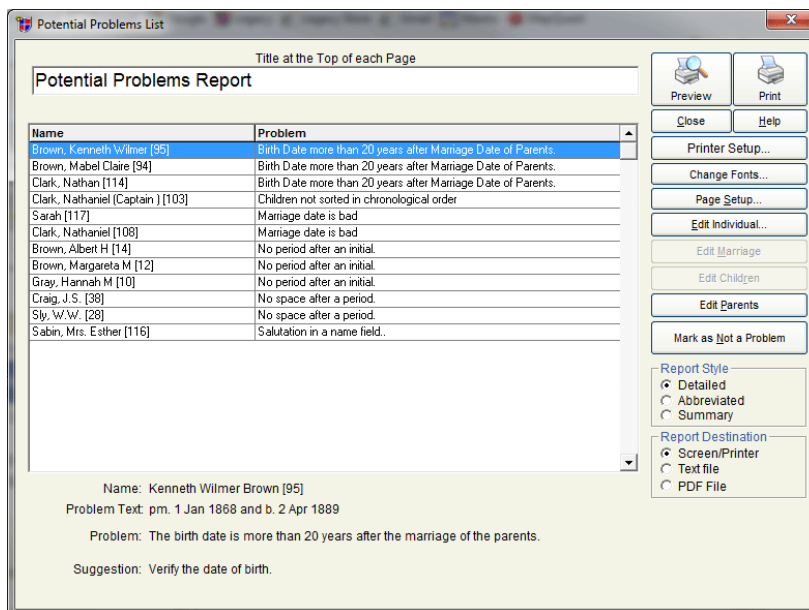
Normal check of all records.

Before clicking on the Create Problems List button in the upper right, look at the Warnings, Problems, Standardization, and LDS (if LDS options are turned on) tabs. Each tab shows you which potential problems Legacy will look at.

Each option is selectable, in other words, if you only want to check your family file for persons who were buried 30 days after their death, select that option, and turn off all the other checkmarks. You can also adjust the values. For example, by default Legacy will warn you if a child was born when the parents were less than 13 years old. You can change this to 10 years old or whatever value makes sense for you.

Now click on the Create Problems List button (no this will not "create" problems, it will create the list of problems).

This image below shows a list of potential problems. The first problem is that Kenneth Brown's birth date is more than 20 years after the marriage date of the parents. Although not the norm, this may very well be true for this child. If so, just click on the Mark as Not a Problem button, and this potential problem will not show up again. But it may be worth looking at, so using the Edit buttons on the right you can investigate further.



Another problem listed is the Nathaniel Clark's children are not sorted in chronological order. This may not seem like a big problem, but clicking on the Edit Children button showed this:

- Asa Clark, born 12 Feb 1742
- Abel Clark, born 24 Jan 1742
- Caleb Clark, born 26 Nov 1744
- ...8 more children

This certainly looks like a potential problem, and before sharing this data with FamilySearch or anywhere else, you should reinvestigate this family.

Narrow your focus

My personal database has information on lots of people that are not closely related to me (8th cousins, 4 times removed, etc.). To begin, I want to focus my clean-up efforts on my direct line.

From the Records tab, I can limit the records being checked to those with a certain "Tag" number. For example, I can mark everyone on my direct line with a tag value of 1. Then I can have the Potential Problems tool only report problems of people on my direct line. For step-by-step instructions on how to use tagging, watch this video: <http://tinyurl.com/y9jhtbx>

Goal: Reduce duplication

One of FamilySearch's goals is to help reduce duplication both in original research and in LDS ordinance submissions. Its new software at <http://new.familysearch.org> incorporates tools to check for duplicates.

Clicking on the new Possible Duplicates tab checks to see if the current individual is listed elsewhere in their databases. The user then has the choice to combine any duplicates.

Before working with new FamilySearch

Before working with FamilySearch, cleaning up and standardizing your Legacy family file is important. Searching your family file for duplicate individuals is an important part of maintaining a well-kept family file. You may have duplicate individuals in your family file for many reasons. Using Legacy's Merge tool, you can easily identify and combine duplicates. Here's how....

1. With your family file open, go to Tools > Merge > Find Duplicates.
2. Now you see the Merge Options screen. Most of the time, just leave everything the way it is and click on the Continue button. If you just want to see a report of potential duplicates, click on the Print button.
3. At the far bottom left of the Merge screen it shows the number of sets of potential duplicates and which number you are currently viewing.

Now, compare the person on the left with the person on the right. The top part of the screen shows the basic information about each person. The bottom half of the screen shows any differences between the two. Using the radio buttons you can reconcile any differences. If you decide that the two individuals are the same, click on the Merge Right Individual into Left Individual button. Legacy then displays the next set of potential duplicates.

If, after comparing the two, you decide that they are NOT the same person, click on the small button in the lower right with the crossed-out circle. This will add this set to the Not Duplicates List, and they will not appear in future Match/Merge sessions.

If you need to pause your merge session because you need to sleep or eat (don't those two things always get in the way of our genealogy time?) just click on the Close button. Legacy will save your place in the merge session. The next time you start Legacy, it will ask you if you want to continue merging.

Now that your family file's duplicates are cleaned up, you are better prepared to work with FamilySearch or any other database.