Genealogy Basics – Ways to Avoid Damage to a Database from a GEDCOM file

By Joe Petrie

In case you don't know what a GEDCOM file is, the following is a quote from Wikipedia:

> "A GEDCOM file is plain text (usually either ANSEL or ASCII) containing genealogical information about individuals, and meta data linking these records together. Most genealogy software supports importing from and/or exporting to GEDCOM format. However, some genealogy software programs incorporate the use of proprietary extensions to the GEDCOM format, which are not always recognized by other genealogy programs, for example the GEDCOM 5.5 EL (Extended Locations) specification."

Unlike Import, Export is easy. Nothing is damaged in the GEDCOM Export file creation process. To protect recipients of my GEDCOM Export files, I select the ASCII option when I produce the GEDCOM Export file. Why? I do not know of any genealogy database software that can't import an ASCII text file.

The producer of the GEDCOM Export file has no role to play in Import – except to give advice. The following paragraphs essentially are my standard advice. First, I recommend that the recipient **not** immediately append any of my GEDCOM Export file's records to their existing database. Instead, I recommend that the recipient Open a new file (named something like JOE) to be the recipient of all the records from my GEDCOM Export file. Why? A cousin did not follow my advice to back-up etc. She imported all of the records from a GEDCOM Export file as soon as she received it. She had to match/merge over 80% of the appended records – because we both had already had the same records. Match/Merge can be tedious!

In addition, before my GEDCOM Export file's records are imported to the new file, I recommend that the user globally add a NOTE to each Individual record (not to the Marriage records) using Word, WordPad, Word Perfect etc. I suggest that the NOTE be an attribution like: GEDCOM from Joe Petrie as of today's date. (Why? That way, all of my errors are mine.)

I also caution the recipient that the modified GEDCOM Export file must be saved as a text file with a TXT file extension – not as Word etc. By definition, the standard format is ASCII text.

Furthermore, I suggest that the new JOE database be loaded by importing all records from the modified GEDCOM Export file. I suggest that the recipient review the JOE database to determine what records are to be exported (to be imported) to the recipient's existing database. I suggest that the recipient first backup the existing genealogy file before a record from their exported GEDCOM file from JOE is imported to their existing database. Why? The back-up is easily restored. (Remember: My cousin had not backed-up her database.)

Finally, each time that a group of records from the JOE file (loaded modified GEDCOM Export file) is added to the recipient's existing database, I recommend that the recipient use the genealogy software's "duplicate" report to identify duplicate records, to be deleted or matched/merged. Yesterday, for example, I matched/merged over 50 records from a partial import of a large group of Mariette records from Sheila Smith's Web Site. I thought that I did a thorough review – not thorough enough! (Sheila's Web Site has a GEDCOM file that I copied.)

#

Email address: joe-apg@norwoodlight.com (your suggestions and comments are appreciated)

Joseph F. Petrie

Founding Member, Friends of Irish Research (FIR)
Member, Cape Breton Genealogy and Historical Association (CBGHA)
Member, Family History Society of Newfoundland and Labrador (FHSNL)
Member, New England Historic Genealogical Society (NEHGS)
Member, Association of Professional Genealogists (APG)