

The Use and Abuses of Online Genealogy

by [Gary B. Hoffman](#)**

People interested in searching their ancestry have never had it easier than today. Why, all they have to do is log on to the Internet, type their name, and their computer will spew forth their entire family tree. Or so goes the urban legend fueled by testimonials of the many neophytes who have done just that. And they have several lines back to Adam, complete with thumbnail photographs. Their genealogy is "done," thanks to the almighty Internet.

Anyone who has pursued genealogy for more than two weeks knows that it's not really that easy.

What is compelling about this scenario is that it is a nirvana we all wish would come to pass. The computer's ease at handling large amounts of information makes it a logical tool (double-meaning intended) for use in compiling and publishing genealogical data. User-friendly programs like Family Tree Maker and the availability of CD-ROMs full of data make it seem easy to just find and download our family data into our personal computer. But like the mirage on the horizon, easy genealogy is forever just beyond our reach.

One technology that makes finding our family data appear easy is the data exchange standard called GEDCOM. Many family historians have compiled their family's data in their favorite genealogy program, saved it in a GEDCOM file, then published it on CD-ROMs, on the World Wide Web, or via the many e-mail mailing lists on the Internet. Others have submitted it to the LDS Church's Ancestral File or their Pedigree Resource File, Genealogy.com, World Connect, and other collections of ancestral data. By contributing to those collections, these family researchers have in essence published their work to the world, making it available to anyone who subscribes to that particular service or downloads it.

Others, usually newcomers to genealogy research, happily search out the names they are looking for, then download all the data they can from these published sources, importing it into their genealogy programs and proclaiming the finished product "their genealogy." Many times, they resubmit this new compilation to the same public outlets and the information returns to the public domain, ready to be downloaded by the next seeker of ancestors.

So what's wrong with this scenario? The problem is that the family information that is freely traded via CD-ROMs and the Internet is often unsubstantiated and of questionable validity. Because it is unvetted, its genealogical value may be very low. Yet people treat it as if it were gospel truth when it should be treated as a clue or a starting point for further research.

In the past, if you read something in the genealogy library or in a family history journal, you could generally count on a publication review process to certify the validity of the data or at least the methodological rigor of the author. Before computers, desktop publishing and the Internet, it was hard to get something into print and wide distribution. Genealogies of reliable authors usually got published only after an editorial review known as vetting. (That's a good word meaning "to subject to expert appraisal or correction"; look it up in the [Merriam-Webster Online](#) dictionary) Even authors who self-published usually took the effort to support their contentions with good source references. After all, they were putting their money on the line as well as their reputations.

Yes, these reliable publications still exist, including many stalwart state and county historical societies and the venerable [New England Historical and Genealogical Society](#), which has published its journal, the New England Historical and Genealogical Register, continuously since 1845.

But the voices of these reliable genealogical sources are being drowned out by the shouts of hundreds of computerized GEDCOM databases on CD-ROM and the Internet. It's so easy to publish good and bad research today; just upload it to a repository. No one checks it and complains if it is deficient. And it is so easy to download, too, much easier than searching out old books in a library. After all, who wants to read a stuffy

academic article with footnotes and alternate theories when a GEDCOM file with (apparently) factual information can be downloaded in seconds?

Yes, that's the dilemma of today's genealogy: that which appears simple is probably unreliable and that which appears equivocal and hesitant is probably more to be trusted. (Of course, the true ancestor detective trusts no source without checking it out.)

So, in the spirit of *caveat lector* (let the reader beware — another good Merriam-Webster term) and *caveat computer* (it could have been good Latin had ancient Rome been computerized), here is my list of do's and don'ts regarding uploading and downloading genealogy from the Internet and CD-ROM repositories.

1. **Know your source.** Who submitted the information? If you can't tell where it came from, how can you check it out? Be sure you notate the file with the source (where you got it) and the submitter. It helps to use commonly accepted citation rules.
2. **Check the documentation.** Are there references for the "facts" given: dates, places, relationships? If so, look them up yourself to check them out. If not, don't give it much credence.
3. **Never import a GEDCOM into your main data file.** Once you add a bogus file into your own data, it's very, very difficult to back it out again. Instead, make a new data file with your genealogy program, then import the data into that first. Look it over and decide if it's worth merging into your good data.
4. **Don't resubmit information that didn't originate with you to a database.** That is, don't recycle someone else's research into your submission without checking all the sources. Especially, don't take credit for someone else's research.
5. **If you do submit your compiled genealogy, be sure you purge the personal information of anyone still living.** Most repositories require you to do this. If you don't, you may be liable for violating those people's privacy and engendering bad feelings in the family.
6. **Use the data from a repository not as authority but as a clue or evidence of a theory.** A GEDCOM file is never a primary source. By definition, it is a compilation of someone else's research or lack thereof.
7. **When you are ready to submit your data to a repository, include references to source materials and lots of justification for the facts in your data file.** Be sure your name and address are associated with the file. Also, be ready to answer queries about the ancestors you are submitting.
8. **Develop a skeptical approach towards downloaded data, just like the hard-bitten detectives in mystery stories.** Make sure each person's birth, death, and marriage information makes sense. Keep asking yourself as you read it, "How do we know this for sure? What proof is there of this?"
9. **Remember, genealogy is a journey, not a destination.** Realize that you'll never be "done" with your quest. Anyway, what's the enjoyment in that? Rather, it's the discovery and learning that we're seeking, not the full filing cabinet.

In genealogy as in other of life's endeavors, the more we learn, the less we know for certain. All the computer has done for us is accelerate the speed at which this comes to pass.

Genealogy.com

****Gary Hoffman**

Gary Hoffman has been involved in genealogy research for over 30 years. He is former president of the Computer Genealogy Society of San Diego and is CGSSD's Webmaster. Currently a computer manager at the University of California's San Diego campus, he recently received a law degree and passed the California bar exam. His articles on technical and legal issues relating to genealogy have appeared in several online publications and newsletters and he is a regular speaker at national genealogy conferences.