Computers vs. Humans: Comparing Output of Genealogy Programs

© 2009, Debbie Parker Wayne

First publication rights granted to *Digital Genealogist, July 2009.*
All other rights reserved.

Computers are very good at handling repetitive, logical tasks—not so good at creativity. Any fan of *Star Trek: The Next Generation* will remember the trouble Commander Data had with tasks that required feelings or creativity. Even the imaginative minds that created *Star Trek: TNG* could not imagine computers had advanced so much by the 24th century that an android could rival the creativity of the human mind.

There has been an ongoing discussion in the genealogical world about whether genealogy programs can "create" output or reports that meet the creative standards we strive for today—more than a list of begats—an interesting family history with analysis of kinship communities and social context that makes the reader care about the family and want to know more. Some genealogists expect a well-written narrative to be magically spit out by a genealogy program's report generator. Genealogy databases allow a genealogical summary report to be generated using canned sentences that Sergeant Joe Friday would love—just the facts. The more advanced programs allow customization of the output sentences and source citations to make the narrative more exciting and interesting. Some genealogists think there is no reason for a genealogy program to offer these options. They want only a bare-bones family summary with birth, marriage, and death information. This summary is copied into a word processor where the sentences that flow from the creative juices are then added. There is nothing wrong with this approach and many of us have worked this way for years.

But what happens when you write the family story as you research, as we are encouraged to do, and your research uncovers a new generation? Or you discover information on a child who was born and died between census enumerations who was never named in any document, until the one you discovered today? Granted, it is easier today to modify the word processor file than it was when you had to use scissors and paste to literally cut and paste pieces of pages together to insert new information. But if the prose is in your genealogy database, a new report can be output with all the people renumbered, in the proper place, and with properly formatted source citations included. If all the creative writing was done in the word processor much of it may have to be recreated or the two documents will have to be manually merged.

We probably all agree that no genealogy program is capable, all alone, straight out of the box, of providing narrative a good editor would think worthy of winning a family history writing contest. But, are modern programs able to store creatively crafted sentences and output those sentences in reports that will require less editing in a word processor than the canned reports will require? That is, if the user can write well crafted prose in a word processor, is it a good use of time to save that prose in the genealogy database instead of a word processor file?

To answer this question we obtained permission to use two articles from the National Genealogical Society *Quarterly* (NGSQ) as project test data. Thomas W. Jones, CG, gave permission to use his article "A Solid Gang of Them"[1] and Elizabeth Shown Mills, CG, gave permission to use her article "Which Marie Louise is Mariotte."[2] Using these articles as test data for the project allowed volunteers to concentrate on the capabilities of their program's report features without worrying about anyone judging their research or writing skills.

As Claire Bettag, CG, a former editor of NGSQ pointed out, the final form of these articles may bear little resemblance to the article originally submitted to NGSQ. The *Quarterly* has a rigorous editorial process where multiple reviewers

---

[1] Thomas W. Jones, "A Solid Gang of Them: An Illinois Morse-Trammell Family's Reactions to Scandal," National Genealogical Society Quarterly vol. 92, no. 2 (June 2004): 105-118.

[2] Elizabeth Shown Mills, "Which Marie Louise is "Mariotte"? Sorting Slaves with Common Names.," National Genealogical Society Quarterly vol. 94, no. 3 (September 2006): 183-204.

and editors suggest changes to the article. No database program or one individual alone can produce such polished output. But if you could write like this, is it possible or advantageous to use your genealogy program instead of a word processor?

Experienced users volunteered to enter the test data and use advanced program features to recreate the articles as closely as possible. The following table summarizes the output from four programs: Genbox, Reunion, Roots Magic, and The Master Genealogist (TMG). Family Tree Maker and Legacy Family Tree volunteers dropped out before completion of this project.

GENEALOGY SOFTWARE OUTPUT COMPARISONS

The native output of each program was compared to the original article. Exact fonts, margin sizes, and other items that vary between publishers were not checked. The use of names and dates, wording of sentences, and inclusion and placement of source citations were compared.

	Genbox	**Reunion**	**Roots Magic**	**TMG**
Output formats	RTF, custom	RTF, TXT, HTML, multiple WP native formats	RTF, PDF	RTF, PDF, TXT, HTML, multiple WP native formats
Supports introduction text and conclusion text	Yes	Yes for introduction, Conclusion entered but it prints before the genealogical summary	Only with Book feature which does not allow embedded source citations	Yes
Allows sentence customization	Yes	Yes	##Yes	Yes
Allows citation customization	Yes	Yes	Yes	Yes
Allows sort date for event order	Yes	Yes	Yes	Yes

Computers vs. Humans: Comparing Output of Genealogy Programs

	Genbox	Reunion	Roots Magic	TMG
Creates true bullet and numbered lists in text entries	No	No	No	No
Child list names in small caps	Yes, but first letter was larger	Yes	No	Yes
Names in boldface or normal weight as in test article	Yes	No	No	No
Sentence fragment to introduce child sections customized as in test article	No	No	No	Yes
Allow different wording in child and adult sections for baptisms excluded from child section as in test article	No	Only names are included in the child list	No	Yes, if baptism is added in the OTHER event category so they are not displayed in the child section with birth data
Child Roman numerals start at "i" for each parent group	No	Yes	Yes	Yes
Generation number included only for the first child in a child list	No	No generations numbers for any child	No generations numbers for any child	Yes
Allow custom paragraphs to precede the child list	No	No	"Spoofed" by adding text to preceding event and reformatting in WP	Yes

Computers vs. Humans: Comparing Output of Genealogy Programs

	Genbox	Reunion	Roots Magic	TMG
Allow custom paragraphs to follow the child list	Yes	Yes	"Spoofed" by adding text to the last child and manually moving it in WP	No, "spoofed" by adding a dummy child, but then the numbering following that child will have to be manually corrected
Compound sentences handled as in test article	No, some printed as separate sentences	Yes	Yes, but placing some information in facts and some in notes to achieve the sentence structure made footnote reference placement incorrect	Yes
Supports indented paragraphs	No	No	No	Yes
Provides footnotes instead of endnotes	Yes	No, but can convert them in a WP	Endnotes in native format, footnotes in RTF format	Endnotes in native format, footnotes in RTF, HTML, and WP formats
Supports embedded citations	Yes	No	No	Yes
Allows citation to a particular data element for an event	Yes	No	No	Can be "spoofed" by adding multiple tags for each data element and combining custom sentences

	Genbox	**Reunion**	**Roots Magic**	**TMG**
Citation elements in same order as in test article (sorted citations)	** No	Yes	Yes	Yes
Citations use ibid or short reference	Ibid.	N/A	N/A	User option, but ibid. not available when using combined citations
Handles *dit* names, accents, umlauts	Yes, in text, not in places where the default name was used such as the parenthetical outlines of ancestry or the first sentence of the genealogical summary	N/A to the test article but software claims to support these	N/A to the test article but software claims to support these	Yes
Miscellaneous Comments	Conclusion was created and then inserted into the report	Periods followed individual names and birth dates even when the sentence continued past the birth date	Cannot handle dates separated by a dash such as 1859—1860 Cannot combine sources or select the order of multiple sources Must remove "XX Generation" headings	Automatically adds a "No known children of XXX" sentence to childless couples unless a tag is added to each couple to suppress the sentence Must remove "Generation XX" headings

WP = Word processor

Compound sentences may require entry into both facts and notes as the text field of a fact is limited to 100 characters.

** This is probably due to the structured output from a genealogy database program and use of freeform entry by the author in the original articles.

CONCLUSIONS

The computer is always going to be better at some things: making sure source citations include all elements correctly in both full and shortened versions, catching typographical errors such as a date of 31 June which should be flagged as a non-standard date, and numbering people consecutively. These elements alone make a compelling case for using a genealogy database for as much of the writing as possible. No matter how many Certified Genealogists and experienced editors review a document, some errors tend to be overlooked by humans.

Experience using advanced program features allows the output to come closer to the desired format. It is important for a program to allow as much customization as possible while supporting the standards of the genealogical discipline for source citations and format. But even the most advanced programs will require some editing in a word processor for the text to meet the standards set forth in *Numbering Your Genealogy: Basic Families, Complex Systems, and International Kin* by Curran, Crane, and Wray. How much editing must be done in a word processor depends on how familiar you are with the advanced features of the genealogy software and how much time you are willing to spend customizing the sentences and sources. If you are writing text that will never be used again, the word processor may be a better use of time. If the text might be reused, spending the time to get it into the genealogy software may save time later. Since no one genealogy program meets every standard in Curran-Crane-Wray or in the *BCG Genealogical Standards Manual*[3] each genealogist needs a checklist of items to change manually in the word processor after a report is generated. For the more sophisticated genealogy programs the list may be shorter. The time spent by the volunteers varied from twenty to fifty hours. More time would have been required if this project had not begun with knowledge of what the output would be. The need to modify repeated location names and the usage of given names and pronouns could be handled in sentence customization during data entry. This would not normally be possible during initial data entry. Space limitations prevent listing every change that was required in a word processor. The table lists only major elements. This was not a perfect test, but may assist each genealogist in determining whether their genealogy program might be used for writing as well as research and analysis.

---

[3] Board for Certification of Genealogists, *The BCG Genealogical Standards Manual*, Orem, Utah: Ancestry for BCG, 2000.

One of the project volunteers stated, "I began this study being somewhat skeptical that genealogy software could be a major component in the process of writing an article. This experience has convinced me otherwise." We all have differing work habits and will have to make our own judgment as to what works for us. Some of us may always be happier using a word processor for writing. But its nice to know some genealogy programs offer sophisticated features that allow incorporating our writing into our database for safekeeping and reuse.