

Ged2Reg User Guide

Draft Version 0.5.3 – 07 April 2021

Ged2Reg is “pre-release” software.

The Ged2Reg software code, design, and this document are protected by copyright.

Copyright © 2020, 2021 by Thomas W. Shanley

At this time: All rights reserved.

Contents

Introduction	3
License.....	3
Conventions, Standards, and Applications	3
User Interface.....	4
So Many Options!.....	4
Large Documents vs. Free Applications	4
Installation	5
Menu.....	6
About “Settings”	6
File Menu	6
Tools Menu	7
Help Menu.....	8
Types of Reports	9
“Main Person” Blocks.....	9
GEDCOM Files	9
Register Report	10
Ancestors Table (Ahnentafel) Report.....	10
Input/Output Tab.....	11
Content Options Tab.....	14
Ancestry Report Tab	19
Citation Options Tab	23
Notes Options Tab	27
Indexes Options Tab.....	28
Rewriters Tab	31
Log Tab	33
Manage Settings Dialog	34
Regex Test Dialog.....	36
Explanation of the Example	37
Supplemental Information.....	39
Styles	39
Marriages vs. Names.....	40
Regular Expressions	41
Miscellaneous Topics	41

Introduction

Ged2Reg (“Gedcom-to-Register”) is a desktop application that can be used to create readable (and editable) text output from a GEDCOM (Genealogical Data Communication) input file. The GEDCOM input is a cryptic data interchange format that can be written and read by many genealogy applications. The output word processing files (docx or odt) present selected content in one of two “standard” formats. The first of these is called Register Format, which was designed for effective communication of lineage (descendants) information to human readers and has long been used in major journals, such as those published by the NEHGS (New England Historic Genealogical Society). The second kind is an Ancestors Table (“Ahnentafel”).

Both of these report conventions are probably familiar to most users who may be interested in Ged2Reg. They are commonly available as options in desktop genealogy applications. Generally, however, those programs treat the report as a final product; further editing may be difficult or effectively impossible. In contrast, Ged2Reg assumes that the user will edit the output in a word processing application, for instance to make minor corrections, to add details not in the input database, to extend the text with hand-crafted family history narrative, or even to paste the entire output into another document as a chapter or section.

License

Ged2Reg is offered for free for non-commercial use, to anyone who can and does accept the license. It is provided with no warranty of any kind. Before you are permitted to use Ged2Reg you must agree to waive any claims to damages of any kind, including without limitation incidental and consequential damages. In addition, you must agree that you are the author of any output that you create using Ged2Reg, and that any possible liability that may be incurred because of what you create is entirely your responsibility.¹ If you do not have the legal capacity to agree to these things, then no license is granted, and you have no right to use Ged2Reg. If you live in a jurisdiction where these terms are not valid or enforceable for any reason and to any degree, then no license is granted, and you have no right to use Ged2Reg. If you intend to use Ged2Reg in any step in an activity for which you expect (or hope) to be paid, such as selling a service or selling genealogy report(s) produced for hire, the free license is not applicable to you, and you need to request terms for a commercial license, and unless and until one is agree you are not permitted to use Ged2Reg.²

Conventions, Standards, and Applications

The content formatting of the Register report is based on conventions in use by the NEHGS since about 1870. A fairly recent book published by the NEHGS³ discusses the conventions and “rules” of this format. The book also discusses conventions applicable to the Ancestor Table format. Ged2Reg itself has many options for formatting its output. The overall structure of Ged2Reg output is based on these

¹ Concerns here include, for example but without limitation, issues such as privacy violation of living persons; or, copyright violations in notes that you choose to include in the output.

² This wording is intended to permit, and the free license is intended to and does allow, an individual working in the context of his/her own family history to use Ged2Reg in producing material that will be, or may be, published in a journal or other periodical, regardless of whether or not some compensation may result.

³ Henry B. Hoff and Penny Stratton, *Guide to Genealogical Writing (third edition)* (NEHGS, 2014):

https://shop.americanancestors.org/products/nehgs-guide-to-genealogical-writing-3rd-edition?_pos=1&_sid=e1c9c3fe2&_ss=r&pass-through=true

conventions, but Ged2Reg has many options that are either outside the conventions or possibly divergent from them in some way. Output that is close to conforming to the conventions is possible; the Ged2Reg UI provides buttons (discussed later) to set, with a single click, the best-conforming choices that it has available.

The output file formats (as distinct from the content) are (supposed to) comply with technical standards for word processing files. For *docx* (Word) files, the applicable standard is *ECMA-376-1:2016 / Office Open XML File Formats* (see <https://www.ecma-international.org/publications-and-standards/standards/ecma-376/>). For *odt* files, the applicable standard is *ISO/IEC 26300-1 / Open Document Format for Office Applications* (see <https://www.iso.org/standard/66363.html>). Fee-based programs such as Microsoft Word 365 (registered trademark(s) of Microsoft Corporation) [aka, Word] and freeware such as Open Office / Libre Office [aka, OLO]⁴ “support” either file format, with greater or lesser completeness and fidelity. Word is primarily used in the development of Ged2Reg. Although either format can be converted to the other by the desktop applications, the conversions are imperfect and it is recommended that you choose the native format for your preferred word processing application. This is especially important with indexes (which convert very poorly between the two) and with footnotes (for instance, if there are “too many” OLO may be unable to open a docx file at all).

User Interface

Ged2Reg is a Windows application that provides a range of options on a tabbed user interface. The bulk of this document is a description of the menu and then of the options, tab-by-tab.

So Many Options!

In its current form, Ged2Reg has a lot of options; probably too many.

Some of them were added on the idea that some element of the report *might* be better done a bit differently from the way it was first implemented; or, it might *not* be better; or, people might *disagree* on which way was better. Others were created in “tension” between different “standards” and/or between the standards and what seems best for a different priority (different from, matching editorial policies of a particular journal) – such as, reducing repetition.

A reasonable way to approach the complexity may be to start with the “Conforming” settings, create some “trial run” report(s) (it usually takes less time for Ged2Reg to *create* a file, than it will take the word processing program to *open* the file), and look at the output. Then look for settings that might have an effect on anything that you see that you might wish was different... such as, say, “stop repeating *United States Federal Census* over and over, please!”.

Large Documents vs. Free Applications

If you expect to be creating large documents (thousands of persons, thousands of pages in one file) and using one of the free word processing applications (Open Office / Libre Office), you should review the information in the *Large documents* and *Working with Libre Office* topics in the section Supplemental Information / Miscellaneous Topics.

⁴ Libre Office was derived from an earlier version of Open Office and has been developed separately. Open Office has released only a few, relatively minor revisions over the past five+ years.

Installation

Download and run the installer.

Ged2Reg now requires the .NET Desktop Runtime 5.0.x to be installed on your PC. This is relatively new and may not already be installed on many systems. In that case, the installer will notify you of the requirement, provide a link to the Microsoft download page, and exit. You will need to download and install the applicable version of the runtime library. You should choose the “latest” 5.0.x that matches your operating system, 32-bit (x86) or 64 bit (x64). After you have completed this step you can run the Ged2Reg installer again and should be able to complete the installation.

After you install Ged2Reg, the program will be available in the Windows menu under

West Leitrim Software > GEDCOM to Register Report

See the Installation Guide for details, if needed.

NOTE: the Microsoft download page for the runtime is

<https://dotnet.microsoft.com/download/dotnet/5.0> ; be sure to download **.NET Desktop Runtime 5.0.x** (currently, **.NET Desktop Runtime 5.0.4**) and not one of the other items with all-too-similar names.

Menu

Most actions are performed using buttons. A few actions are implemented in the menu. Most of these pertain to “settings”.

About “Settings”

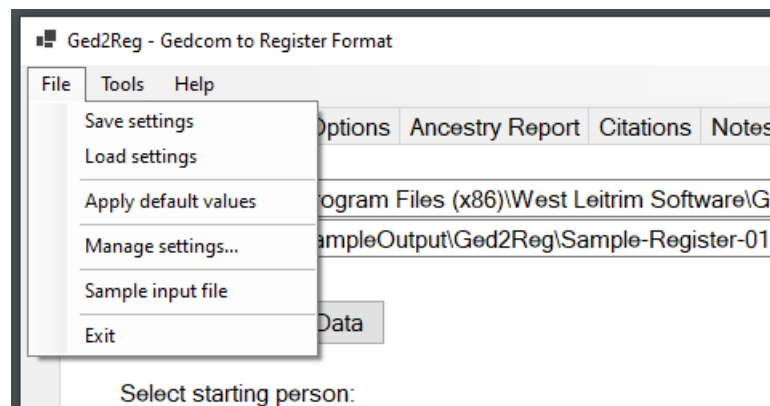
When launched Ged2Reg automatically loads the previously saved settings, or default settings if none were saved. Ged2Reg will also ask if you want to save settings when you close it.

Ged2Reg saves settings (when you request it) in a single file in its own new directory under the Windows-defined application data directory for the logged-in user. When you start, there is one “set of settings” in the file, and its name is “DefaultSettings”. The DefaultSettings set is created automatically when you first start Ged2Reg. You can immediately start changing settings to suit your preferences, and you can save the settings using the menu or in response to the prompt when you exit Ged2Reg. Then, when you start Ged2Reg again later, the saved settings are automatically loaded and you can proceed to work with them, or change them, and so on.

That design is intended to make things as simple as possible for “casual users”. A more “advanced” (more complicated) way of working with settings is also available. Suppose, for example, that you expect to provide different “flavors” of output, say one that is “complete” with citations, notes, etc. and one “streamlined” with options adjusted to reduce the number of pages the reports take up. You may then wish to maintain multiple sets of settings, with different names that are meaningful to you. When Ged2Reg is running you can easily switch between the different sets of settings that you have saved.

File Menu

All of the settings actions on this menu apply to all of the settings on all of the tabs, there is no need to repeat the menu actions on different tabs.



Save settings – saves all “sets of settings” to disk, including the one you are currently using and any others that you have previously created (and not deleted).

Load settings – reads in the file with all “sets of settings” that you have saved, and selects the DefaultSettings as current. Any unsaved changes to settings (current set or otherwise) are lost.

Apply default values – resets (most of) the current settings to their original values (as if you had first started the program). The set name and current file names are retained automatically, and you are asked if you want to retain any title rewriters.

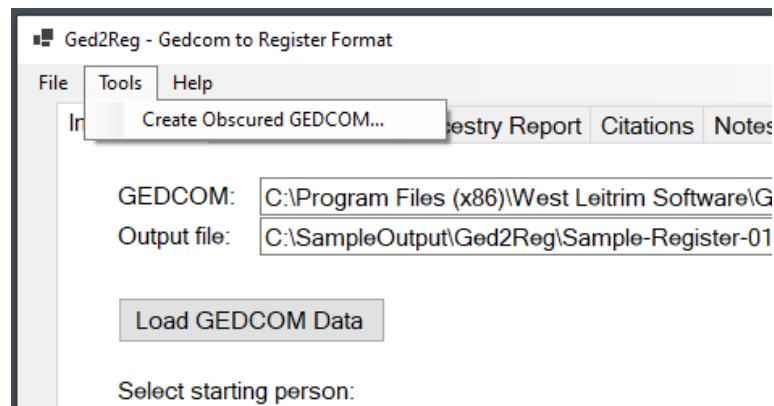
Manage settings... – this opens a dialog window where you can work with named sets of settings that you create; this is described in the section *Manage Settings Dialog* (“casual users” may prefer to skip it).

Sample input file – this option sets the name of the input file (GEDCOM text box) to a sample file that is included with Ged2Reg, and loads it. Also, a “starting person” is pre-selected, based on your current settings: a person with plenty of ancestors if you are set for an Ancestors report, otherwise a person with plenty of descendants. Also, if there are no “title rewriters” in your current settings set, a few rewriters that may be applicable to content in the sample file are added.

You might wish to use this sample data in running some “test” reports to verify that Ged2Reg is working and to get a look at some expected output with various options. You are, of course, free to choose different starting persons.

Exit – prompts to ask if you want to save settings, then closes Ged2Reg.

Tools Menu



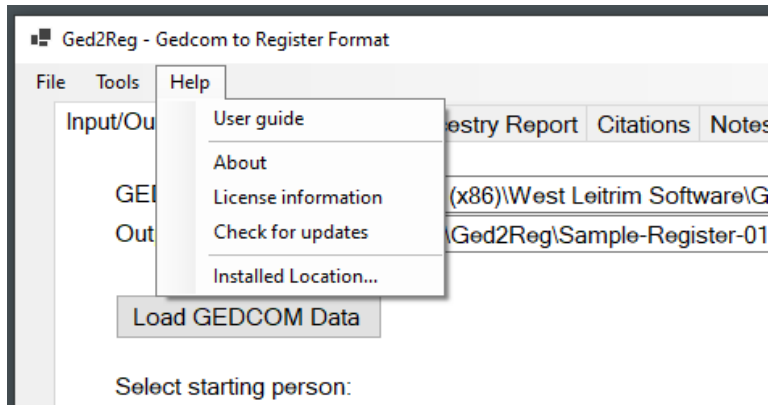
Create Obscured GEDCOM... - this option lets you write out an “obfuscated” (“obscured”) copy of your current input GEDCOM file. The process replaces given names and surnames with randomly selected ones from pools of hundreds of more-or-less common names; and, changes exact dates by a few days (plus or minus 30, randomly selected for each date); and, truncates all web links that it recognizes to just the host name; and, changes all exact street addresses that it recognizes as such to randomly assigned street numbers and names; and, removes “page” information from all citations; and, drops all “repository” and “object” (media file) information. Name and date changes are not applied to the records for any person that the process recognizes as having been alive at any time prior to the year 1750.

A final in-memory scan is made of the complete text of the file, and if occurrences are still found of the most-common surnames in the original input file they may be “fixed” in a final pass. Because of potential damage to the GEDCOM structure in this process, a second “alt” output file is created in this final pass. Details are written to the Log tab.

This feature was developed for, and used in, creating the sample file that is embedded in the Ged2Reg installation. If you, a Ged2Reg user, find some reason to use it, please do be aware that it is **not** “guaranteed” to eliminate all possibility of identifying real persons named in a GEDCOM. For example, event descriptions or notes may have been typed in with variant (or incorrect) spellings of names, and those will typically not be recognized and will not be changed. Also, since some names are also commonly used words, the name-replacement process may occasionally turn a note or other text into a

puzzler, say by changing “Stake and Stones” into “Stake and Armstrongs” (assuming the random reassignment of the name Stone in this case was to “Armstrong”). Also, the date-change logic can result in odd-looking cases, such as person being baptized a few weeks before they were born.

Help Menu



User guide – opens (this) user guide document in your default application for pdf files.

About – displays a simple dialog with the name and (build) version number. The first two parts of the build number match the “release number”, the last two parts are unique to the build (compile).

License information – displays again, the terms that you agreed to in order to use Ged2Reg.

Check for updates – opens the “home” of this program on github, in your default web browser. There you can look for a newer “release” on the mid-to-lower right-hand side of the page.

Installed Location – displays the directory where Ged2Reg is installed. Resources such as the sample GEDCOM file are stored in or below that directory.

Types of Reports

Ged2Reg now provides a choice of two different kinds of output reports: a descendants report (Register format) and an ancestors report (Ahnentafel). The choice is controlled by an option on the “Ancestors report” tab, which is described in a later section.

Each report starts with a “root” person whom you select, but they work from there in opposite directions, one going “up” through the starting person’s ancestors, and the other working “down” through his or her descendants. The numbering systems are completely different. And, obviously, the two make different choices regarding the persons to be included in the output. But, within those different contexts, the detailed content for each person who is included in the output will be similar.

“Main Person” Blocks

In several places in this document we will refer to a “main person”. This is not the same as the “starting person”. There is only one “starting person” but there will be many “main person” blocks in your report.

When we refer to a “main person block” we are discussing the sequence of one or more paragraphs starting with a number at the left margin that is immediately followed by the person’s name in bold caps, and extending through the listing of the children, if any (and if the report includes them). A “main person” is the one whose name appears at the start of the block.

The starting person always appears as the first main person.

GEDCOM Files

GEDCOM is a file format that was devised decades ago by genealogy software developers to serve as a mechanism for transferring data between different genealogy applications. It is meant to be read mostly by software, not so much by people. Still, it is a text format that can be opened with text editor application such as Notepad.



```
sample.ged - Notepad
File Edit Format View Help
0 @I2393@ INDI
1 NAME Eugene /Adams/
2 SOUR @S1155@
3 PAGE Page info [removed]
3 DATA
4 TEXT Births, Marriages, Deaths from 1800 (earliest entries evidently written
5 CONC in 1830 by Eugene Adams)
2 SOUR @S1021@
3 PAGE Page info [removed]
3 DATA
4 TEXT Record for Willm Adams
3 NOTE @N569@
3 _LINK http://search.ancestry.com
2 SOUR @S1016@
3 PAGE Page info [removed]
3 DATA
```

In some circumstances you might wish to look at your GEDCOM file. In that case you might want to enable the *Include debug output* setting, so that your report will include identifiers that you could then search for in the file.

There are many online references to GEDCOM. Wikipedia may be a good starting point:

<https://en.wikipedia.org/wiki/GEDCOM> .

Register Report

This kind of report starts with a selected individual, to whom the number 1 is assigned. After one or more paragraphs about the starting person and spouse(s), the children are listed. Children who had children of their own are assigned sequential numbers, 2, 3, and so on. The report then continues through generations of descendants of the one starting person.

1. **EUGENE W¹ ADAMS** was born at Beckley, Sussex, England about April 1769,¹ baptized at Beckley on 29 April 1769,² and he died at Rye, Sussex, England on 18 May 1848 (age 79).³ He was buried at Brede, Sussex, England on 11 April 1848.⁴ He married first Rhonda Patton at Brede parish church, Brede, Sussex, England on 28 July 1795 ...|

Children of Eugene W¹ and Rhonda (Patton) Adams:

- i. **LILLIAN² ADAMS**, b. at Brede on 17 November 1797,⁵ bp. at Brede on 03 January 1798.⁶
2. ii. **EUGENE ADAMS**, b. at Brede 21 February 1800; m. first Anna Hansen, second Delia Swanson.
- iii. **GENE ADAMS** b. at Brede about February 1801⁷ bp. at Brede on 10 March 1801⁸

A good starting point if you want to know more about the format is American Ancestors, the web site of the NEHGS, which first developed the Register format about 150 years ago. See for instance <https://www.americanancestors.org/Browse/Publications/The-Register/Writing-a-Family-Sketch-Register-Style> .

Ancestors Table (Ahnentafel) Report

This kind of report also starts with a selected individual to whom the number 1 is assigned. After a paragraph about the starting person, the parents are listed. A person's father is assigned the number that is double that of the person's own number, and the mother is assigned the number double plus 1. An equivalent paragraph is presented about each of the parents, then their parents are included in the same way, proceeding through the ancestors of the one starting person, and potentially listing as many as twice as many people for each successive generation.

1. **WALTER YATES**, son of Max Yates and Alberta Kirk, was born at Brookfield, Fairfield County, Connecticut on 10 November 1877,¹ and he died at Danbury, Fairfield County, Connecticut on 19 November 1922. He was buried at Danbury on 30 December 1922 (Wooster Cemetery).² He married Cheryl Stark, as her first husband, at Danbury on 17 April 1896 (by Rev E. L. Whitcome).³ She was the daughter of Nicolas and Teresa (Randall) Stark, born at Danbury on 20 July 1876,⁴ and she died at Bethel, Fairfield County, Connecticut on 29 April 1943 (122 East Street).⁵

2. **MAX YATES**, son of Geoffrey Yates and Vivian Savage, was born at Danbury, Fairfield County, Connecticut on 16 January 1852 (said to have been "originally from Roxbury"),⁶ and he died at Brookfield, Fairfield County, Connecticut on 10 October 1878. He married Alberta Kirk, as her first husband, at Brookfield on 16 December 1876.⁷ She was the daughter of Joshua and Bernadette (Warner) Kirk, born at Brookfield on 06 September 1857,⁸ and she died at Danbury on 08 August 1941 (Danbury Hospital).⁹

3. **ALBERTA KIRK**, daughter of Joshua Kirk and Bernadette Warner, was born at Brookfield, Fairfield County, Connecticut on 06 September 1857¹⁰ and she died at Danbury, Fairfield

If you want to know more about this kind of report, Wikipedia may be a helpful starting point; see <https://en.wikipedia.org/wiki/Ahnentafel> .

Input/Output Tab

Input/Output
Content Options
Ancestry Report
Citations
Notes
Indexes
Title Rewrites
Log

GEDCOM: C:\Program Files (x86)\West Leirim Software\Ged2Reg\Resources\sample.ged
Output file: C:\SampleOutput\Ged2Reg\Sample-Register-01.docx

Load GEDCOM Data

Select starting person:

Name	Life Span	Place Born	Place Died
Briggs, Edgar	1769 - 1848	Beckley, Sussex, England	Rye, Sussex, En
Briggs, Edgar	1863 - 1863		
Briggs, Edward	1850 - 0000	England	
Briggs, Edward	1877 - 0000	Heathfield, Sussex, England	
Briggs, Eileen	1958 - 0000	Danbury, Fairfield, Connect...	
Briggs, Ellen	1930 - 0000	Bethel, Fairfield, Connecticut	
Briggs, Ellis	1927 - 0000	New York	
Briggs, Elmer	1873 - 1948	Ulster, New York	Queens, New Yc
Briggs, Elsie	1909 - 0000	New York	
Briggs, Elvira	1839 - 0000	Ashburnham, Sussex, Engl...	
Briggs, Emily		Heathfield, Sussex, England	

Generations to report: 5
Open After Create ☒

Create Report
Cancel

Current settings set:
Sample--Descendants

GEDCOM – the input file. Names end with “.ged”. The name can be typed, or the button can be used to open a standard *Open File* dialog. If the dialog is used, the file is opened after (and if) you click OK.

Output file – the full path of the output file. Names end with “.docx” (to create Word files) or “.odt” (to create Open Office files). The [...] button can be used to invoke a standard *Save File* dialog, where you can locate or create the output directory and enter the file name. Nothing happens to the file until and unless you later “Create Report” (see below).

Load GEDCOM Data – clicking this button opens the input file and links up all the data in memory so that the report can be produced. You would use it if you typed or pasted in the file name, and/or to open the file after you launch Ged2Reg or switch between settings sets. It may take a little while, e.g., perhaps 30 seconds or so for 50,000 persons. When the data have all been loaded some key information for all the individuals included in the GEDCOM file is listed in the table occupying most of this tab. As you may notice in the screen shot, unknown years are filled in here with “0000”. This is done only to align the information in the column in this display; the 0000 years appear nowhere else.

Select starting person – after the input file is opened, all the individuals in it are listed in the grid. Initially it is sorted by Name. You can click on the column headers to sort by other columns. You can use the keyboard to jump to the first row with the last name starting with a given letter. Use the scroll bar to navigate the list. If you expand the overall Ged2Reg window, the grid will get bigger along with it.

Click anywhere in a row to select the starting person. For a register (descendants) report, this will be earliest ancestor in your report, from whom the others are descended. For an ancestors report, this will be the latest descendant.

If you create a report and then save settings, and then later re-open the same file (with the same settings set selected), the starting person you previously picked should be selected automatically. (The starting person is remembered separately in each settings set.)

Generations to report – number of generations of descendants to include in the output. The starting person is “1”. An additional generation (i.e., one more than the number chosen here) may appear in limited fashion if the nth generation has children. Technical note: there are quirks in the processing of these “up/down” fields: “sometimes” the changes you key in may fail to “stick”. To work around this, it is suggested that you either use the arrows to change their values, or press Enter after typing a value, and switch to another tab and back to verify the value you want has been retained. “Sorry about that”.

Open After Create – if checked, after a successful create (see below) the output file will be opened with your system’s default application for the file type. If you do not have a word-processing application (Word, Open Office, Libre Office, ...) set up as the default application, Windows will usually default to using Wordpad. That built-in application “can” open the files, but it does not support some essential features of either format. This checkbox remains enabled while a report is being created, so you can change your mind in-process. The state of this checkbox is not considered a “setting” and is not saved, it always starts out as “off” when you launch Ged2Reg.

Create Report – this button is disabled until you load GEDCOM data. When clicked, Ged2Reg disables all changes to settings “for the duration” of the report processing and enables the Cancel button. While processing the report Ged2Reg performs further in-memory manipulations of the GEDCOM input data, based on the options selected, and then emits the report into the output file. If a file of the same name (and location) already exists you will be asked if you want to proceed by deleting the old file.

This process can take some time if there are many generations of descendants (or ancestors) to process. The time required can vary by options selected, e.g., turning off citations altogether may be much faster. For example, a test case that includes about 5000 “main person” descendants and uses the most-complicated options for selecting citations runs about two minutes. Progress messages are displayed on the status line at the bottom of the window and are also added to the Log tab. When finished a summary of the process is also output to the Log tab. The summary includes counts of the numbers persons of various categories that were included in the output.

Cancel – this is enabled while the processing is in progress. To stop processing, click this button. At the next opportunity (usually a few seconds) Ged2Reg will stop processing and re-enable the settings. No output is produced; there may be an empty file, and if you chose to over-write an existing file, the previous version will have been deleted.

Current settings set – this field displays the name of the settings that you are currently using. If you have not created multiple, named sets of settings (which is optional, and done using the *File > Manage*

settings... menu item) it will always display the “DefaultSettings”. If you have created multiple settings sets, you can use the drop-down list in this field to quickly switch between them. Switching sets changes the options on all tabs to the values in the chosen set and clears any GEDCOM data that was previously loaded and displayed in the table.

The current values of the settings set that you switch away from are remembered, and will be available if you switch back, and will be saved if you save the settings.

Content Options Tab

Settings on this tab control the content of a descendants (Register) report and, where applicable, also control aspects of the content of an ancestors report.

Within the body of the Register report, each person with descendants (excluding the starting person) will appear twice, once as child of their parents and once in a “main person block”. Here these may be referred to as “continued children”. Those without descendants will appear only once, as a child of their parents.

A main person block may contain B-B-M/D-D-B (“vital events”: birth, baptism, marriage/divorce, death, burial) information; child entries usually contain somewhat less, especially continued children (because it is all repeated later). The information that is output can be influenced to some extent by option settings.

Input/Output	Content Options	Ancestry Report	Citations	Notes	Indexes	Rewriters	Log
Title	<input type="text" value="Sample Report"/>						
Author	<input type="text" value="Ged2Reg User"/>						
When to include baptism in output	<input type="button" value="Always"/> ▾	Italicize names in lineage list <input checked="" type="checkbox"/>					
Include (event descriptions) / edited	<input checked="" type="checkbox"/> <input type="checkbox"/>	Reduce output for continued children <input type="checkbox"/>					
Include Burial Events	<input checked="" type="checkbox"/>	Minimize output for continued children <input type="checkbox"/>					
Omit Burial Date	<input type="checkbox"/>	'Standard' output for children / abbrev. <input checked="" type="checkbox"/> <input checked="" type="checkbox"/>					
Generation numbers on all children	<input type="checkbox"/>	Shift all-caps names to mixed case <input type="checkbox"/>					
Obscure (possibly) living names	<input checked="" type="checkbox"/>	Reformat unknown names <input checked="" type="checkbox"/>					
Omit living persons' citations	<input type="checkbox"/>	Unknown names in the input <input type="text" value="_"/>					
Age limit when guessing	<input type="button" value="99"/>	Unknown names in the output <input type="text" value=""/>					
Reduce place names when repeated	<input checked="" type="checkbox"/>	Include Report Summary <input checked="" type="checkbox"/>					
Full place name once per generation	<input type="checkbox"/>	Include Debug Output <input type="checkbox"/>					
Drop 'USA' from place names	<input checked="" type="checkbox"/>	<input type="button" value="Conform settings"/>					
Inject word 'County' into place names	<input checked="" type="checkbox"/>						
Place before Date	<input checked="" type="checkbox"/>						
Starter generation 'number' (optional)	<input type="text" value=""/>						
<i>Most recent use of this settings set</i>							
Date and time	<input type="text" value="3/16/2021 9:48 AM"/>						
Processing time	<input type="text" value="00:00:00.8416334"/>						
Output file created	<input type="text" value="C:\SampleOutput\Ged2Reg\Sample-Register-01.docx"/> <input style="float: right;" type="button" value="..."/>						

Title, Author – when entered, these will be set as properties in the output file. Typically, if you later use your word processing application to create a PDF file these properties will be carried forward to it.

When to include baptism in output – the choices are Always, When No Birth, and Never.

Include (event descriptions) / edited – each fact in a GEDCOM file may include a date, place, and a short text description. This option controls whether or not the descriptions are included in the output. Ged2Reg makes some attempts to prevent this text from damaging the sentence structure of the output: the text of the description may be modified, and it is always placed in parentheses. The second checkbox, designated “edited”, influences the degree of modification of the text. If the option is on (checked), then Ged2Reg will scan the whole text the description field and, for instance, convert any periods (“.”) that it thinks are sentence terminators (as distinct from, say, abbreviations) into semicolons. If this option (second checkbox) is “off” then Ged2Reg will only examine the ends of the description text, and convert leading uppercase characters to lower case (unless the first word appears to be an abbreviation, part of a name, or part of a title), and drop any trailing period.

Burial Events – you can choose to include or omit them. If included, you can optionally omit the date.

Generation numbers on all children – different publishers disagree on the question of whether the generation number should be shown on the first child in a set of them (NEHGS), or on all children (NYGB). This option lets you choose based on your preference.

Obscure (possibly) living names – you can optionally have Ged2Reg “guess” if the individuals selected for the report may still be living, and in that case to replace the person’s given name with “(Living)”. If there is a death event for the person (with or without any date or place) they will be assumed to be deceased. Otherwise Ged2Reg will guess how old the person would have to be, to be still living, based on the available information including the person’s birth or marriage dates, spouse’s birth or death date, children’s birth dates, etc. **NB: enabling this setting does not guarantee that a reader will be unable to determine the real-world identity of living person(s) that are included in your data and fall within the scope of the report.** You, the user, are the author of the output report and you are solely responsible for your content and any consequences of the information you include.

Omit living persons’ citations – if this is also checked, then source citations will not be emitted for facts about persons whom Ged2Reg guesses may be living. Since the citations may often mention the person’s name, this setting may help increase the effectiveness of “obscuring” them.

Age limit when guessing – this number represents the “maximum plausible age” used in guessing if a person may still be living.

Reducing the length of place names. Place names, when fully spelled out, can take up a lot of space and reduce readability. But if they are never fully spelled out they may be ambiguous. Several options control how Ged2Reg handles place names in place name fields (i.e., not in description fields, and not in notes).

One way to reduce the length of place names is not to repeat the county and state. If both of these options are selected, the second one is honored; if both are “off” then this kind of reduction of place names is not applied:

Reduce place names when repeated – This option, if selected, directs Ged2Reg to shrink the place name, if possible, after the first occurrence within each main person block. For example, a three-part name consisting of Town, County, State will usually be reduced to the town. This option represents the “standard” way of handling place names and is recommended.

Full place name once per generation – This option can be set to provide that Ged2Reg will start over with the full place names beginning at each generation instead of at each main block. This option is “non-standard” but may reduce “clutter” and improve readability in some cases.

Drop ‘USA’ from place names – It is usually clear when a place name refers to a location in the USA, but genealogy software often appends the country name regardless, and if output to your report this can be very distracting. If this option is selected, the various ways of representing USA will (always) be omitted from the place names associated with facts included in your report.

Inject word “county” into place names – The word county is typically omitted by genealogy applications but is generally expected to be included, where applicable, in “standard” Register style. Select this option to have Ged2Reg add the word “County” to that part of a place name, when it recognizes that that is what it represents and it doesn’t already have a similar designation (such as Parish). Since two-part place names are ambiguous (is the first part the county, or a town?) they are not processed; in most cases, the county name in a three- or four-part place name in the USA, UK, and Ireland will be recognized. Note: Ged2Reg does check these names against an internal list of known county names, and will skip unknown “middle names” of places, which might be mis-spelled county names, or might be “something else”. Enabling this will also enable expansion of the abbreviation “Co.” into “County”.

Place before Date – the convention is somewhat loose but does call for “place” before “date” when giving the details of a vital event. To “conform”, this option should be on (checked). If you prefer dates first, uncheck this option.

A note about “at” and “in” and “on”. The convention is that an event occurs “at” a place if the specific location (e.g. town) is known, unless it is a major city. But events occur “in” a major city, or “in” a country, state, or county if the place is only known to that level. Ged2Reg uses a short list of 10 major cities to “nod” to this convention. You may wish to make edits in your output if you prefer this be expressed differently. As for dates, events occur “on” a specific day but “in” a month or year.

Starter generation “number” – there is a distinction, in the numbering convention in the case where the line of descent being reported included both an immigrant and one or more of his/her ancestors. The rule for this case is, the immigrant is generation “1”, the immigrant’s parents are generation “A”, the immigrant’s grandparents are generation “B”, and so on go backwards through the alphabet as you go “up” to earlier generations. If this situation applies to the report you are creating and you wish to have this convention applied, you should fill this field with the letter that applies to the first generation (starting person). So, for example, if the starting person you have selected was the grandfather of the first immigrant, you would enter “B”.

Otherwise you should typically leave this field blank. You could put in a number greater than 1, and that will be applied in the generation numbering, but the starting person will still be number one.

Italicize names in lineage list – this refers to the list of names of preceding generations that follow the name of the main person in each main person block. “Standards” disagree as to whether or not these names should be in italics; you can set this option on or off to match your requirements or preference.

Detail for ‘continued children’. You have some control over the details output for a child who also later appears as a main person. If you select none of these options then all the available / applicable information will be output. If you select more than one of them, then the last one (in the order listed here) “wins”.

Reduce output for continued children – select this option to reduce repetition by including fewer details in the listing of child who appears later in a main block.

Minimize output for continued children – reduce continued child entry to just the name and lifespan years.

Standard output for children / abbrev. – if this is selected, then the information shown for each child is given in a terse style that is based on the *Register* “standard”, as we understand it. For a continued child this consists of the name, birth date and place if known, with baptism as a substitute if not, and spouses’ name(s) if known; citations are deferred to the later appearance as a main person. If the second (abbrev.) box is also checked, the event words such as “born”, “baptized”, and “married” are abbreviated as “b.”, “bp.”, and “m.”, respectively.

This two-part setting also controls non-continued children, who will be listed in a similar fashion but with additional events (since they appear nowhere else).

Shift all-caps names to mixed case – if your data includes personal names in ALL CAPS, it’s possible that you want it that way; but it does interact poorly with the conforming style of “small caps” for main person and child names. If this option is on, Ged2Reg will shift all names that are entirely in upper case in the input, to mixed case in the output. Some of the obvious special conditions – “Mc” names, “de la” names – should be handled correctly. Still, with this option enabled, you should be attentive to the possibility of needing to edit the output to clean up some converted names.

Reformat unknown names – select this to change the way unknown names are output.

Unknown names in the input – tells Ged2Reg how to recognize unknown names in your data. Common representations include underscore, dash, question mark, but this is not limited to a single character.

Unknown names in the output – how names that are recognized as “unknown” (based on the setting above) should appear in the output. The standard representation (and default) is a blank space shown as a sequence of five underline _____ characters.

Include Report Summary – if this is checked then a page is appended to the end of the report, listing most of the options used, recording the date and time, and giving counts of the number of main persons, citations, etc.

Include debug output – emits some diagnostic information (principally, the GEDCOM ids of persons and sources) into the output report. For technical use only: the output will not be usable for any other purpose.

Conform settings (button) – click this button to change the settings in the current set (all tabs) to best approach conforming with the conventions as we understand them and as Ged2Reg features allow. This button appears in two places, here and on the *Ancestry Report* tab. Clicking this one will also turn off the Ancestry report option, if it is set.

Most recent use of this settings set – these display-only fields show information about the last time this settings set was used to create a report – when it was completed, how long it took to run, and what file was created. PLEASE NOTE some limitations on this information! It is only available if the settings were saved after the report was created. And, it does not provide assurance that the settings are the same now as they were then.

[...] (**button**) – if there is a file name displayed, clicking this button will attempt to open it, using your default application for the file type. If the file no longer exists (or, has been moved) a message will be displayed instead.

Ancestry Report Tab

Settings on this tab allow you to select and configure the optional ancestry report (Ahnentafel) output. A few of the options, as noted, also have an effect on the descendants report output.

Input/Output	Content Options	Ancestry Report	Citations	Notes	Indexes	Rewriters	Log
Options unique to Ancestry report:		Options that also affect Register report:					
Output Ancestors Report	<input checked="" type="checkbox"/>	Generation headings	<input type="checkbox"/>				
Omit generation superscripts	<input checked="" type="checkbox"/>	Reduced margins	<input checked="" type="checkbox"/>				
Allow multiple appearances	<input checked="" type="checkbox"/>	Use host name for link text	<input checked="" type="checkbox"/>				
Placeholders for unknowns	<input checked="" type="checkbox"/>						
Space between couples	<input checked="" type="checkbox"/>						
Include back references	<input type="checkbox"/>						
Also list siblings	<input type="checkbox"/>						
Generation prefix numbers	<input checked="" type="checkbox"/>						
Focus on one ancestor (see below*)	<input checked="" type="checkbox"/>						
Omit focus spouse(s)	<input type="checkbox"/>						
Continue past focal ancestor	<input type="checkbox"/>						
'Merge' duplicate ancestors	<input type="checkbox"/>						
Minimize from generation	<input type="text" value="0"/>						
Also drop back-references	<input type="checkbox"/>						
*Focused ancestor							
Currently selected ancestor	<input type="text" value="Charlemagne (died 0814) [7]"/>						
<input type="button" value="List choices"/>	<input type="button" value="Apply selection:"/>	<input type="text" value="Charlemagne (died 0814) [7]"/>					
<input type="button" value="Conform settings"/>							

Output Ancestors Report – select this option to create an ancestry report instead of a Register report.

Note: for a “conforming” report (i.e., one that best matches the usual conventions for this kind of report), the next four options in this column should be “on” and the rest of them “off”. That is the case after clicking the “Conform settings” button (see below).

Omit generation superscripts – select this option to omit (not output) generation number superscripts on the main person names in the ancestry report. This (omit them) is the usual convention.

Allow multiple appearances – select this option to allow an ancestor from whom the starting person descends in multiple ways, to appear multiple times in the output. This is the usual convention. The second (and subsequent) appearance of a person in the report does not include the person’s details. Instead, it just lists the name and the number of the first appearance.

If this is “off”, then an ancestor will only appear once in the output, the first time that person appears in the tree. This may be useful if you are interested in reducing the size and amount of repetition in the report.

Placeholders for unknowns – select this option to output placeholder lines, with the appropriate ancestor number, for ancestors who are not known (within the input GEDCOM). This should be ‘on’ for a conformant report. Where there are multiple consecutive unknowns, the placeholder is written as a range. Turning this option on makes the report longer, but it can never more than double the size of the report. If you prefer, you can also turn placeholders off in distant generations by setting “Minimize from generation” (see below).

Space between couples – convention suggests making the report easier to read by inserting blank space between couples to make them easier to see. Check this option to conform to that.

Include back references – if this option is selected, a child block is output after the parent(s). This is not the most conventional way to structure this kind of report, but it may be helpful to a reader trying to navigate the family. The child block is similar to that output in the Register report format.

Also list siblings – if this option is selected *in addition to* “Include back references”, each person’s sibling(s) are also listed at the point where the child entry is inserted. Again, this is “non-standard” but may be helpful in orienting the reader, or in presenting the case on a particular line, etc.

Generation prefix numbers - this option will add the generation number, followed by a dash character, at the start of the person number wherever it appears in the report. If your report has many generations this may be helpful in navigating the report, in discussing a particular entry, etc.

Focus on one ancestor (see below*) - this option, in all its variations, is a departure⁵ from the usual conventions in this kind of report. When you select this option (and select a specific ancestor of the starting person, see below), the only individuals listed in the report *between* the starting person and that “in-focus” ancestor will be the ones who are on the direct line(s) from the starting person to that ancestor.

One reason you might wish to do this, is in the case where you have identified a particularly significant ancestor, say, one with proven royal ancestry. Running a report with that ancestor selected “in focus” will isolate the lines of descent that lead from the interesting ancestor to your starting person. You might use this to help concentrate your research, for example, to further study and support (or disprove!) any weak points in the lineage. You might use the focused report to communicate your findings to your family and friends. And, if this focused study leads you to realize that one line to that ancestor is better-supported by evidence than some other(s), you could choose to emphasize the stronger line.

Each generation between the starting person and the focused ancestor may be represented in the report by only one person, or, if you choose not to ‘Omit focus spouses’, by as few as two. But if there is more than one path through the tree from the starting person to the focused ancestor⁶ there will be a

⁵ So far as we know, this option is original and unique to Ged2Reg.

⁶ It is very common for persons with ancestry in 17th- and 18th-century New England to find that they descend in multiple ways from some of the same ancestor(s), making them, in some degree, cousins of themselves.

corresponding (though generally small) increase in the size of those intermediate generations. When the focused ancestor does occur multiple times in the starting person's ancestry, the "focus" logic will continue to apply (and keep the report "narrow") up through the last (oldest) generation where the selected, "in focus" ancestor appears.

If this option is off, then the other settings below that relate to it are ignored (except "minimize", as mentioned); so, they do not need to be individually cleared to stop their effects.

Omit focus spouse(s) – a focused report will, by default, include the spouses of the in-line ancestors. This option, if selected, will omit them from the report (they are still mentioned in the text of the main paragraph for the in-line person).

Regardless of the setting of this option, and unlike a usual Ahenetafel report, the report while in-focus does not expand to include spouses' ancestors, unless the spouse is also a descendant of the focused ancestor.

Continue past focal ancestor – by default, the focused report will stop at the last (oldest) generation in which the focused ancestor appears. If this option is selected, the report will instead continue from that point, with a normal handling of the ancestry.

'Merge' duplicate ancestors - this option may be useful if your tree was assembled by merging multiple sources and has "duplicate" entries of "the same" person. When this option is "on", the "focus" logic will look for and group individuals with the same name and lifespan, and treat them as one and the same, only for the purposes of the "focused" report.

Minimize from generation – if enabled, this option will reduce the details listed for each person to a minimum from the indicated generation (1 being the most recent, i.e., the starting person). This may be helpful in long reports where there are many repetitions of "ancient" ancestors.

Also drop back-references - this option, if enabled in addition to "minimize...", will drop back references starting from the same point.

Currently selected ancestor – this field shows the name and lifespan for the currently selected, "in focus" ancestor. A number in brackets may also appear; this represents the number of distinct lines of descent from the focused ancestor to the starting person; seven lines, in the example, from Charlemagne down to the sample GEDCOM's default starting person for ancestry reports.

List Choices – use this button to analyze the tree and find all ancestors of the starting person. This process includes "merging" duplicates if that option is selected, and it also detects and counts the number of different places that each person appears as an ancestor of the starting person.

When completed this will fill in the selection list in the drop-down box toward the right; that list is otherwise empty, since Ged2Reg does not do all that extra work unless you ask it to.

Apply Selection – after you List Choices and then select an entry in the drop down list, click this button to accept the choice and remember it along with the settings. This is what actually makes a particular ancestor "in focus".

Note that if you complete this action and later change the starting person (on the *Input/Output* tab) the selected ancestor may or may not appear in the starting person's tree. The selected ancestor also will

not appear in the tree if the number of generations chosen is too small to reach that ancestor. In either case an error message will be displayed after you click *Create Report*.

Conform settings – this has the same effect as the button with the same label on the *Content Options* tab, except that it turns on the *Output Ancestors Report* option.

The following options also affect the Register report (they are included here for technical reasons).

Generation headings – select this option to insert headings between generations in the output report.

Reduced margins – select this option to reduce the (rather generous) default margins to a uniform one inch on each side. Margins, and other settings, can also be changed later in your word processing application.

Use host name for link text – if this option is off, hyperlinks that appear in footnotes will be shown as the full text of the link. If on, the link will remain the same internally but the displayed text will be reduced to the “host name”. There are multiple factors to consider here. If your distribution will be on paper, the reduced links will not be useful, since e.g. all of your links to images on Family Search will look exactly the same. If your distribution is primarily electronic (usually, pdf file) then the question is, are most of your links meaningful to a human reader? For example, Wikipedia links are often readable, whereas Ancestry.com links are usually a long string of gobbledygook that conveys nothing to a human.

Citation Options Tab

Input/Output

Content Options

Ancestry Report

Citations

Notes

Indexes

Rewriters

Log

Citation selection strategy
PriorityDriven

Citation fill-in strategy
LeastOftenUsed

Output as end notes instead of footnotes
☐

Summarize additional citations
☒

Max number in summary
2

Repeated citations:

☐ Use 'See Note #'
☐ Use note number inline
☒ Defer/combine consecutive repeats

Use brackets around footnote numbers
☐

Omit footnotes on child that is continued
☒

Placeholders for missing citations
☒

Format for 'See Note #' citations (use '{0}' in Format to insert value)

	Seq	GEDCOM Tag	Format	After
▶	1	Source TITL	{0}	
*				

Full citations (change Seq and sort by column to re-order)

Prefer edited citations
☒

	Seq	Name	Format	After
▶	1	Source AUTH	{0}	,
	2	Source TITL	{0}	:
	3	Citation PAGE	{0}	,
	4	Citation DATA.TEXT	{0}	

Priority patterns (highest 1st, one / line)

certificate
will
probate
(deed)|(mortgage)

Negative priority patterns

public member tree
(one world tree)|(OneWorldTree)

Source citations are critically important and dauntingly complex. An important issue is the fact that users of today's desktop genealogy applications are apt to accumulated many "noise level" source citations for certain facts. For instance, one person's birth event might cite ten or more census records (5 or 6 for the person themselves giving age and/year of birth, place born, etc.), and then any number for the person's children (giving the parents' place of birth). Collecting all this is *good practice* for the family researcher but *reporting it all is very rarely useful* or necessary; rather, that would amount to a cluttered file dump reducing readability. But neither GEDCOM nor desktop applications have a widely used mechanism to indicate that a particular source citation is the preferred one.⁷

The overall strategy used in Ged2Reg is to emit as a footnote or endnote (at most) one source citation for each fact, with an optional summary of other known sources appended to it. To further reduce

⁷ The Quality field, GEDCOM QUAY data element, could be used for this purpose, but the author has not seen this done in any way that seems effective.

clutter, when the same source citation is chosen for successive facts for a single person, it can optionally be emitted only once, after the last fact in the sequence.

Limits may apply to the total number of footnotes. At this time it is known that docx is limited to 32,767 footnotes; Word cannot open and cannot repair any file with more than that number footnotes. When a recognized limit is reached, Ged2Reg will stop emitting footnotes and start counting the ones that cannot be emitted. If that count is non-zero (displayed in the log, and listed on the optional stats page at the end of the output file) you may want to take corrective action. One option is to reduce the number of footnotes. Some settings on this tab, if turned on (checked), may help reduce the number of footnotes; especially consider the settings *Defer/combine consecutive repeats*, *Use note number inline*, and *Omit footnotes on child that is continued*. Other settings here, if on (checked) may work to increase the number of footnotes; especially consider *Placeholders for missing citations*. If you cannot reduce the number of citations below the limit, your only options may be to reduce the number of generations (or, accept the resulting output).

Settings on this tab control aspects of source citation processing within Ged2Reg.

Citation selection strategy – if the input file contains exactly one source citation for a particular fact, then Ged2Reg will always show it (unless this option is set to None). This setting controls how Ged2Reg chooses the source citation to use, when a fact has *more than one* attached to it. The choices are:

- **Priority Driven** – First, pick the source citations that are likely to be “best”. This option requires both “priority patterns” (see below) and a fill-in strategy (without which, many facts may remain uncited).
- **Most Often Used** – Pick the source citations that are used on the largest number of facts that will appear in the scope of the particular report. This will result in the fewest number of distinct source citations, considered over the span of the entire report.
- **Least Often Used** – Pick the source citations that are used on the smallest number of facts. You might choose this on the idea that the sources used less widely are more specific to the facts they document, and so, perhaps, they are usually better choices.
- **None** – select this to turn off citations altogether.

Citation fill-in strategy - this setting is only meaningful if the Priority Driven strategy was chosen. It controls how Ged2Reg chooses the source citation to use for all the facts that have multiple source citations in your data but for which none was chosen based on the set priorities (e.g., suppose that your priorities are “certificate” and “will”, but the fact cites only a couple of census records). The choices are as above, but “None” and “Priority Driven” here will result in no further citations being chosen, and the output will include only the citations that match your priorities (which may be a good thing, if that’s the way you want it).

The fill-in process considers source citations that were already chosen for other reasons, including (1) based on a priority pattern, and (2) on account of being the only citation for a given event; so that if they are linked to other facts for which a citation was not yet chosen, they may be preferred. Also, during the fill-in process, the sense of “most/least often used” is in reference to events that do not yet have a source citation chosen for them, as the program logic goes through multiple “rounds” of filling in citations.

Output as end notes instead of footnotes – if selected, notes come at the end instead of the bottom of the page. By default within Word the end notes numbering style is Roman Numerals; that gets ridiculous with thousands of notes, and so, Ged2Reg changes this setting within each output file. If you really wanted them, you can change the numbering style in your word processing application. Note: throughout this discussion we will refer to “footnotes”; except where explicitly stated, those mentions apply equally to “endnotes”.

Summarize additional citations – if selected, when a fact has multiple citations Ged2Reg will append a short summary (by Title) of as many as **Max number in summary**, and then finish with a count if there are still more.

Repeated citations – the default, if all of these are “off”, is to output a separate, complete footnote for each fact as it appears in the output document. You can choose one of these options if you prefer not to have the second and subsequent use of the same exact source citation repeated in full. Be aware of interactions among them, if you choose more than one of these options: “use inline” overrides “see note”; “defer/combine” overrides “use inline” but combines with “see note”.

Use ‘See Note #’ – select this option to output a brief note referencing the first occurrence of a particular citation, e.g. when multiple family members cite the exact same census record.

Use note number inline – select this option to emit the same note number again. This probably yields the greatest reduction in repetition and in the size of the output. On the down side, it can make it more difficult for a reader who actually wants to check your citations

Defer/combine consecutive repeats – if enabled, wherever the same citation has been selected for consecutive facts on the same person, it will be footnoted only once, after the last in the sequence. And when that happens (i.e., the same footnote is being used to support more than one fact), the footnote text will say so, by appending a short statement such as “(Cited for: Birth, Baptism.)”. You could somewhat increase the impact of this choice by also choosing a citation selection strategy of “MostOftenUsed” (although, that may not be the choice that provides the most compelling evidence).

Use brackets around footnote numbers – footnote numbers and generation numbers look the same (superscripts) and may appear confusing. Select this option to surround the note number with [square brackets] so they look different from generation numbers. NB: this may require additional work if you intend to add more footnotes during editing in your word processing application, since none of the desktop programs will create consistent (i.e., bracketed) numbers when you insert more footnotes.

Omit footnotes on child that is continued – select this option to reduce redundancy by omitting the citations on the listing as a child of a person appearing later in the report as a main person (where all the citations will then appear). Note: a choice of *Standard output for children* on the *Content Options* tab also has this effect; in which case the setting of this option here is irrelevant.

Placeholders for missing citations – if this option is checked, in addition to emitting footnotes based on your other settings, Ged2Reg will output a placeholder footnote, inline, for every fact that has no source citations in the input file. The text of the footnote will be an “eyecatcher” such as “***Citation needed*** for: Baptism”. You could use this option in an early “draft” and go back to your genealogy application to add some evidence, and/or edit the output document to fill in or delete the placeholder footnotes as seems best.

Priority patterns – this only applies (and is required) if you select Priority Driven as your strategy. You may enter one or more keywords, one per line, to help identify preferred sources. Upper- and lower-case letters are treated the same in this processing. The entries are ordered, such that one appearing earlier will be chosen in preference over those appearing later. Note that these patterns are actually regular expressions, and if you know what that means you can try entering patterns more complicated than simple keywords.

Negative priority patterns – these are optional, and similar to the above, but used to down-rate matching source citations when your strategy is something *other than* Priority Driven. For example, if you commonly have facts that cite *Public Member Trees* and something else, say census records, you should include Public Member Trees here to increase the likelihood that the less-derived source citation (i.e., something other than a public tree) will be chosen for the report.

Format for ‘See Note #’ citations – this controls which information about the source is emitted before the ‘See Note #’ link. The {0} is required in order to actually output the content.

Format for full citations - this controls which information about the source is emitted in a footnote. The pieces are assembled in the sequence that they are listed; you can re-order them by editing the Seq cells and sorting on that column (clicking the column heading). If your source citations seem to be incomplete you may want to look into how your genealogy software outputs the information into GEDCOM and choose different pieces (or more pieces: you can add rows by typing into the empty one at the end) by making different selection(s) in the Name column. The {0} is required to output the chosen field; you should put any punctuation in the After column, as that helps the program to omit the entire piece when the field is empty for a given source citation.

The most likely case for you to make changes here: if you don’t like wording that Ged2Reg emits around URLs, this is where you can change that.

Prefer edited citations – in FTM (Family Tree Maker [trademark of Software Mackiev]), in addition to entering source citation information into separate fields for author, title, etc., a user can optionally edit the assembled citation text in any way that they may choose. When an FTM user does this, FTM not only stores the edited citation internally, it also emits the edited version into the GEDCOM in a custom tag (_FOOT). If you want to use these edited citations, where they exist, instead of having Ged2Reg assemble the citation text, turn this setting on (checked). Assembly of the citation text will proceed ‘as usual’ for any source citation that does not include this custom tag in the GEDCOM; for any that do include it, the edited text will be used, **and** any rewriters that would otherwise apply are bypassed.

Also please note that the citation choices made by Ged2Reg are not the last word for what will be in your final output. If you don’t like the chosen citation: edit the content and change it. If you think some particular facts really need multiple source citations to back them up: edit the document and insert another footnote. And so on; you are the author!

Notes Options Tab

Input/Output	Content Options	Ancestry Report	Citations	Notes	Indexes	Rewriters	Log
<div> <div> Include Each Main Person's Notes <input checked="" type="checkbox"/> </div> <div> Include Main Person Spouses' Notes <input checked="" type="checkbox"/> </div> <div> Note Intro Line Format <div>Additional information for {0}:</div> </div> <div> <div>Bold text</div> <input checked="" type="checkbox"/> </div> <div> <div><i>Italic text</i></div> <input checked="" type="checkbox"/> </div> <div> Convert text that looks like divider lines <input checked="" type="checkbox"/> </div> <div> Set notes off with divider lines <input checked="" type="checkbox"/> </div> </div>							

Include Each Main Person's Notes – if you have person-level (not fact-level) notes in your data and wish to have that text emitted into your report select this option. The notes will follow the main person block for each person (not the listing as a child). Note: as the word “Each” should remind you, a “main person” is not the same as the “starting person”: there is only one starting person, but there will be many “main persons” in your report.

Include Main Person Spouses' Notes – same as above, but for spouses; this is independent of the Main Person option.

Note Intro Line Format – text used to introduce notes. Optionally include {0} to mark the place to output the person's name; recommended especially if both (person and spouses') are enabled. You can optionally choose to have this text emitted as **Bold text** and/or *Italic text*.

Convert text that looks like divider lines – if enabled, Ged2Reg will look for lines in the notes that consist of entirely dash characters or entirely equals characters and output them as either a single or double horizontal line, respectively.

Set notes off with divider lines – if this option is checked, each person's notes that are output will be set off by a single horizontal line at the start and end.

For any person with no person-level notes in the GEDCOM file, the other elements (intro, dividers) will not be output.

Indexes Options Tab

	Name index	Place index
Enabled	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Columns	<input type="text" value="2"/>	<input type="text" value="2"/>
Heading	<input type="text" value="Index of Names"/>	<input type="text" value="Index of Places"/>
Index name	<input type="text" value="names"/>	<input type="text" value="places"/>
Separator*	<input type="text"/>	<input type="text"/>
Reset	<input type="button" value="Default"/>	<input type="button" value="Default"/>
+ married names	<input checked="" type="checkbox"/>	
- continued child	<input type="checkbox"/>	

* Use TAB to right justify page numbers (this is the default)

Ged2Reg can optionally mark-up the output it produces with the “hidden fields” markup needed support automatically created index(es) of names and/or places.

Indexes are the aspect of the report that converts most poorly between docx and odt files, when you use one of the desktop applications to open a file the other one’s native format. The two file formats do not provide for identical index capabilities. And neither program fully understands the other’s index options. So, if you do intend to use the indexing, you should not expect to switch between desktop programs during editing; create the output for the program you (most) intend to use, and stick with that one when working on the file.

The settings shown in the screen shot are intended for a docx file, where the result will be to produce two separate indexes. If you use the same settings but create an odt file, you will get one index encompassing both names and places (intermingled), with the first heading. If you use these settings to create a docx file and then open the file with Open Office or Libre Office, the index will be unusable trash.

The settings apply independently to the two indexes (docx) or to the two kinds of entries (odt):

Enabled – if (and only if) this is checked, Ged2Reg will emit (hidden) index entry fields throughout the text at each appearance of a personal name or a place name, and an Index field at the end of the output. Reduction of place names does not impact the indexing. Any similar content (text that looks like personal or place names) that may appear in notes, will **not** be indexed.

Columns – the number of columns for the word processing program to use in formatting the index; values are restricted to either 1 or 2.

Heading – this text is emitted at the end immediately before the index field (place where the actual index will appear). In odt files, even if both indexes are enabled and both have a Heading value, only the first will appear in the output.

Index name – (MS Word / docs only) to keep the indexes separate if both indexes are enabled, each must have a different, non-blank name; this is used in the markup to separate the two kinds. Has no effect in odt output; any value here results in a mess if used in docx files later opened in OLO.

Please note that naming the index complicates the process of adding additional, by-hand index entries in Word; each index must include the appropriate name, but the dialog does not give you a place to enter it directly. If you are going to index content that you add to the document in Word, you may find it works out better if you enable only one index or allow them combine into one by clearing the *Index Name* for both.

Separator – used to tell Word how to format the index; a TAB character (not visible here) tells MS Word to right align them; this field does nothing in OLO.

Reset – click the button to restore default settings for the given index.

+ married names – enable this option to add (presumed) married names of women to the index.

If this option is off, a woman who was born Mary Jones and later married twice, once to a Smith and once to an Adams, and whose name you might thus type, in its final form, as “Mary (Jones) (Smith) Adams” will be listed in the index ONCE, as “Mary Smith”. Each page that she appears on, married at that point or not, will be included in that index entry.

Turning this option on does not change that “Mary Smith” index entry; it adds additional entries in her (presumed) married name(s). In this case there would be two additional entries, one for “Mary (Jones) Smith” (listed with the Smiths, and only showing the page numbers where she is listed as married to Smith) and one for “Mary (Jones) (Smith) Adams” (listed with the Adamses, and only showing the page numbers where she is listed as married to Adams). This may seem complicated when described, but in practice it is not so difficult and the index with this convention applied is clear enough.

PLEASE NOTE: *because of limitations in the GECOM data standard, so far as we understand it, this logic requires an assumption of a certain convention, that of a woman taking a husband’s surname at marriage. We recognize and respect the fact that it has become increasingly popular, from say the latter half of the 20th century, for a woman to make a different choice as to her name when married. But, this assumption still has to be made here, as an unavoidable aspect of programming this feature using the available data. This is not meant to disparage or to disrespect any persons choosing to name themselves in ways that are not correctly predicted by assuming the convention. Users who may take offense, or are concerned about a potential to give offense, by creating index entries that might misrepresent a woman’s name should either not enable this option or else edit any unwanted index entries out of the text.*

- continued child – if your report is large and you are using OLO, you may find that the word processing application is unable to build the index. For example testing shows that “sometimes” Libre Office 6 simply hangs when asked to update the index in a file containing 25,000 to 35,000, or more, index entries. If you have a problem of this sort, you may need to forgo indexing altogether, but might first try turning off *married names*; and you might also try enabling this option, which will further reduce the number of index entries by skipping the places where a continued child appears in the document as a child.

Important note about indexing and page numbers. Page numbers are unknown to Ged2Reg; they do not exist until the word processing application performs page layout. So, the indexes can be *configured* in the output document but they cannot be *filled in* by Ged2Reg; indexes are located at the end of the file but will be initially filled with only a simple line of placeholder text in place of all the index data. To populate the index, you need to open the file in your word processing application and scroll or jump to the end. There you can right-click on the placeholder text for each index, and select the appropriate action, e.g., “Update Index” (“Update Field”). This may take some time to complete as the word processing program will need to repaginate the entire file to perform this action.

Rewriters Tab

Name fixer (regex / replace expression)

(?-i) ((([] (Mrs.) | (\d[a-z]{2})) . *? []))

Event description fixer (regex / replace expression)

(?i) (?<part1>^.*?\d{4}) (?<part2>[a-z]{2,}.*\$)

\$_{part1} \$_{part2}

	Applies	Skip First	Replace Entire	Input	Output
	Everywhere ▾	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Kirchenbuch [Churchbook], Katholi...	Kirchenbuch,
	Everywhere ▾	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	genealogy of the Quick family in A...	Quick Genealogy
	Everywhere ▾	<input type="checkbox"/>	<input type="checkbox"/>	United States Federal Census	US Census
▶	Everywhere ▾	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	The Whitney Family of Connecticut	Whitney Family
*	▾	<input type="checkbox"/>	<input type="checkbox"/>		

<
>

Special purpose “fixers”. The top part of this tab presents two special purpose “fixers”, one that applies to names and one that applies to event descriptions. Each consists of two parts – a “regex” (regular expression) that works as a finder, and replacement pattern that is applied when and where the finder succeeds.

Each of these also comes with a button; if clicked, the button opens a dialog that provides a simple (i.e., minimally helpful) way to test the finder/fixer against some sample text that you provide.

The subject of regular expressions is quite technical and complicated. Most users will not have occasion to use this feature of Ged2Reg, or may forgo using it due to the level of difficulty. Some further discussion is provided later, in the discussion of the “tester” dialog. But, if the subject is unfamiliar and you wish to possibly use these features, you will need to seek out some resources outside the scope of Ged2Reg.

Name fixer – these fields let you define an optional finder/fixer that applies to personal names. Processing is applied separately to the given names and surname subfields. A match results in the entire matched name part (given names OR surname) being replaced using the pattern (second field).

The example in the screen shot was developed in conjunction with a particular GEDCOM file. The data included many cases of placeholder names for spouses (wives) whose names were unknown, in two forms: (*Mrs. ...*) or (*1st Mrs.*) (parentheses included). The finder/fixer combination shown converts these names into conforming unknown name placeholders.

Also, note that this processing applies, and may alter names, *before* indexing markup is constructed.

Event description fixer – these fields let you define an optional finder/fixer that applies to event descriptions. Processing applies separately to each event (fact) that has text in the description subfield. The example shown in the screen shot is discussed later, in the section on the *Regex Test Dialog*.

Title Rewriters. Genealogy reference works often have very long titles. While it may seem necessary to cite the full title at least once, citing it over and over again can take up a lot of space, crowding content off the page and making the report harder to work with. In the table of Title Rewrites you can control how and when source titles are shrunk to improve your output.

In the grid you can easily add new entries by typing into the empty line at the bottom. You can delete lines by clicking on the “selector” on the left side of the row and hitting the Del key on your keyboard (though there is an alternative way to disable an entry, see below).

The values in each row are:

Applies – what part of the output to apply this rule (line) to. Choices are:

- Nowhere – use to turn a rewrite off without deleting it.
- Full Citation – applies in the complete listing of the citation.
- See Note – applies when emitting a See Note reference to an earlier citation.
- Others List – applies within the “Other sources include...” text.
- Everywhere – all (3) of the above

Skip First – do not apply the reduction to the first occurrence.

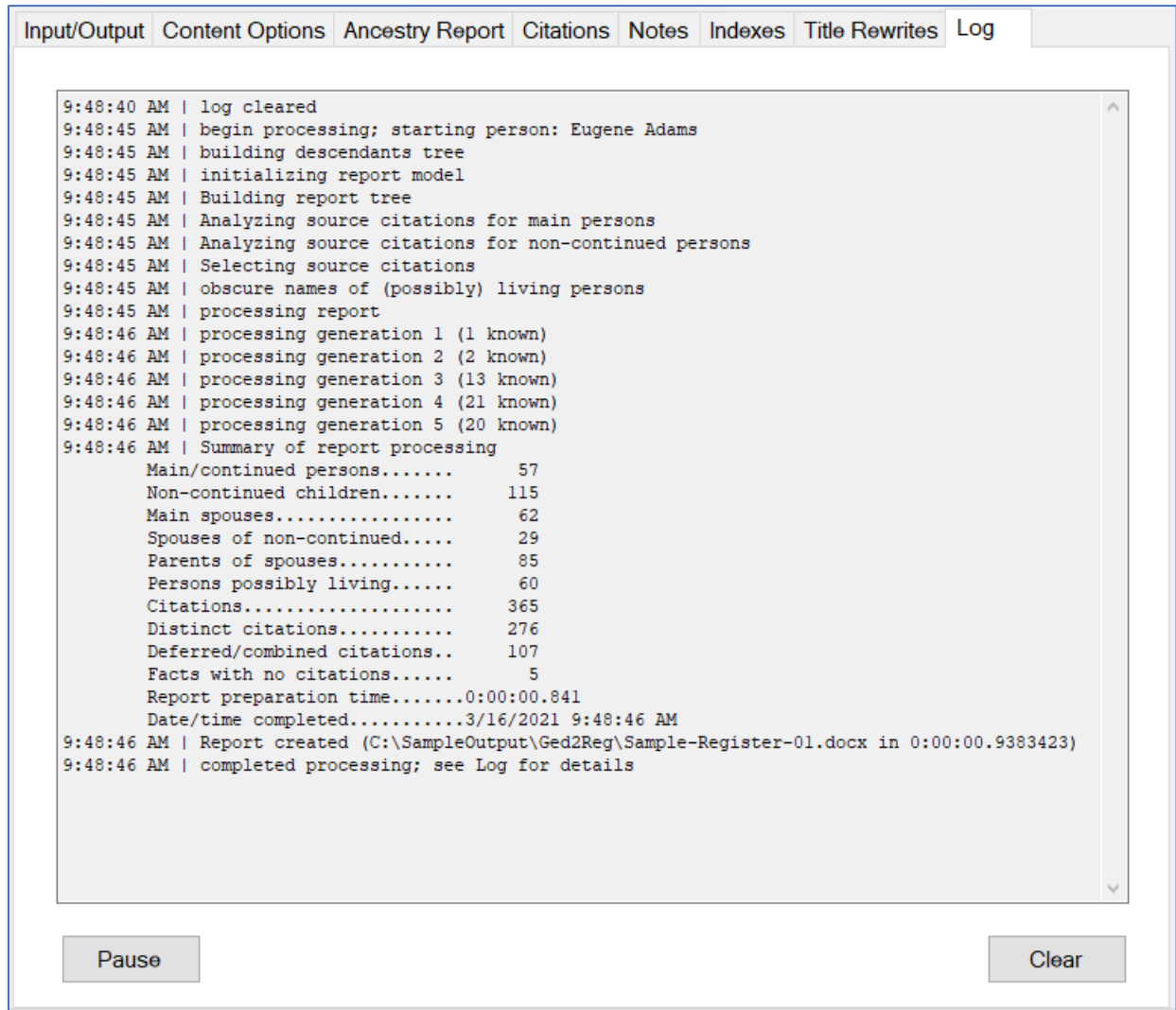
Replace Entire – if checked, the entire matching title is replaced. Otherwise, only the exact part of the title text that matches the input is replaced. The example for census clarifies this: with this option off, that rewriter will transform “1810 United States Federal Census” into “1810 US Census”, and transform “1820 United States Federal Census” into “1820 US Census”, and so on. Also, the example input for *Whitney Family* is not the complete title, but it matches enough of the actual title to be unambiguous and is set to replace the entire title.

Input – this is the text that is searched for in the title. Note that if you are copying and pasting this from a trial-run Word document, it is likely that an extra space will have been tacked on at the end, whether you want it or not. Spaces are significant, and may prevent recognition of a match, so, be sure to check.

Output – this is the text that is output in the case of a match (when other conditions are also satisfied).

Log Tab

This tab holds a running record (in-memory only, not saved to disk) of status and informational messages that may be displayed by Ged2Reg. Also, if any unexpected exceptions (technical errors) occur, information about the exception will be listed here.



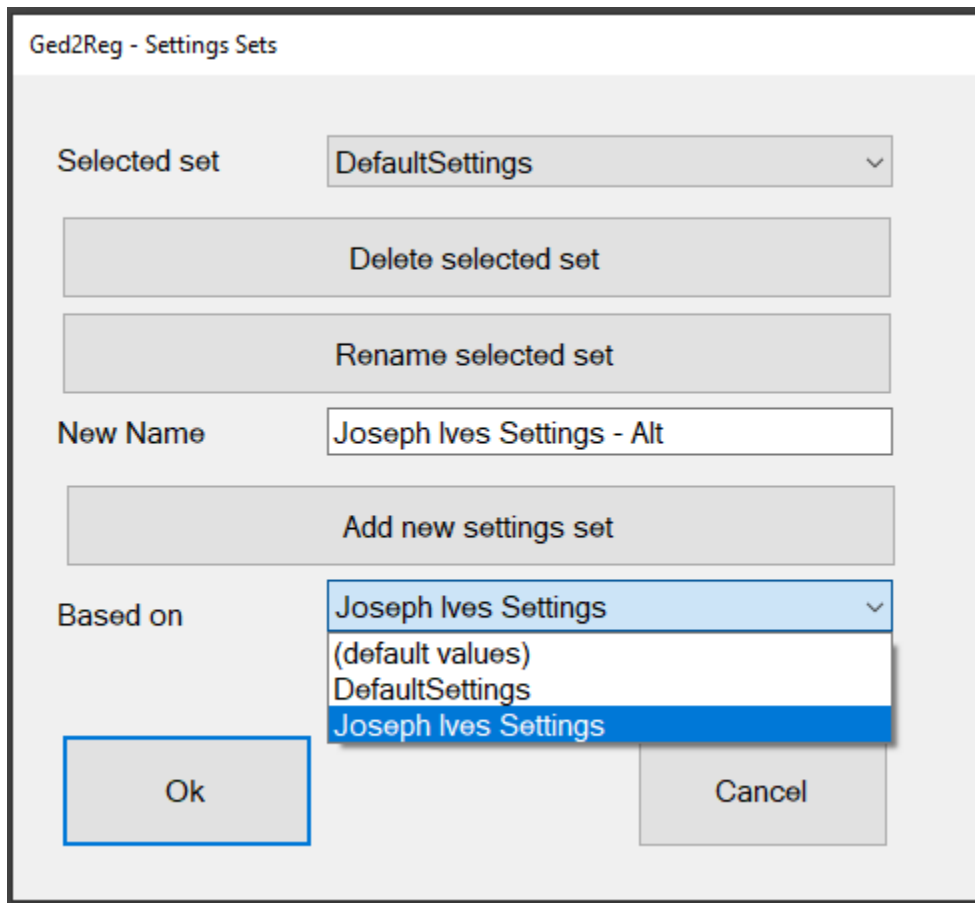
(Text box) – this read-only field displays the messages. If you want to retain a copy of information displayed here, you can right-click to Select the text and to Copy it to the clipboard, then paste the text into some other application (e.g., email, Notepad, etc.).

Pause – this will suspend output of new messages to the log text box, which may be helpful if you want to stop it scrolling. When you click Pause, it changes to Resume. If processing is in progress it will continue, and messages may stack up to be displayed when you click Resume.

Clear – the contents of the log text box will be deleted.

Manage Settings Dialog

This popup window is accessed by selecting the *Manage settings...* action on the *File* menu. It is entirely optional, and only needed if you prefer the flexibility of quickly switching among multiple sets of settings. While this window is open the main Ged2Reg window is locked; to return to it you must click either *Ok* or *Cancel*.



Selected set – the name of the settings set that is currently selected for Ged2Reg to use. When open the dialog after starting ged2Reg, it will show the name of the settings set you are currently working with on the main Ged2Reg window. The simplest thing you can do on this window (other than cancel) is to click on the drop-down and pick a different set, then click OK. That is another way switch between sets (the dropdown list on the *Input/Output* tab lets you do the same switching, a bit more quickly and easily).

The first time you access this dialog after installing Ged2Reg there is only one settings set available, the DefaultSettings. You will need to add more, using this dialog, if you want to be able to switch.

Delete selected set – this prompts with a warning, and if you proceed, the selected set is removed, and the selection reverts to the default set (until and unless you pick a different one). An error message is displayed if you try to delete the default set.

Rename selected set – this prompts with a warning, and if you proceed, applies the text in the **New Name** field to the selected set, and refreshes the lists so that the new name appears. This action is not “reversible” in that later hitting Cancel will not undo it. An error message is displayed if the **New Name**

is blank or if there is already a set with that name. An error message is displayed if you try to rename the default set.

Add new settings set – this button is the only way to create additional sets of settings. The new set will have the name in the **New Name** box. It will be initialized with settings from the currently displayed choice in the **Based on** dropdown list; you can pick “(default values)” or any of the settings sets you have already created, to copy them over as a starting point. This is the situation in the screen shot: we are ready to create an alternative copy of the “Joseph Ives Settings” that we could then edit (on the main window) to produce some different result or to try out some options without the risk of losing track of the way it was working before. After the new set is created the lists are refreshed and the new set is the selected one. An error message is displayed if the **New Name** is blank or if there is already a set with that name.

Ok – any and all changes you made are passed “in memory” back to the main Geg2Reg window, and the **Selected Set** is applied. Changes that you then make to options on the various tabs of the main window will affect the selected set and not any other one. To keep the changes for another day, you will need to save settings either using the menu or at exit.

Cancel – changes that you made (other than renaming a set that already existed) are discarded. The options throughout the tab pages of the main window are not changed.

If you type something into the *New Name* box and later click either button without taking any action that used the new name (i.e., add or rename), a message is displayed to ask if you are really finished.

Regex Test Dialog

Two buttons on the *Rewriters* tab provide access to this optional dialog, which can be used to try out a regular expression rewriter on some sample text, which you might copy-and-paste from a trial run.

The screenshot shows a dialog box titled "Ged2Reg - Regex Tester". It has a light gray background and standard window controls (minimize, maximize, close) in the top right corner. The dialog is divided into several sections:

- Rewriting Event Descriptions**: A section header.
- Sample text**: A text input field containing the text "in about 1790-1795something went wrong".
- Recognizer pattern (regex)**: A text input field containing the regex "(?i) (?<part1>^.*?\d{4}) (?<part2>[a-z]{2,}.*\$)".
- Replacement pattern**: A text input field containing the replacement pattern "\${part1} \${part2}".
- Test / results:**: A button with a blue border.
- Results area**: A large text area displaying the results of the regex test. The text is as follows:

```
Construct regex with pattern [(?i) (?<part1>^.*?\d{4}) (?<part2>[a-z]{2,}.*$)]
Matching text [in about 1790-1795something went wrong]...
...True
Groups in match:
...Entire match, 'in about 1790-1795something went wrong'
...part1, 'in about 1790-1795'
...part2, 'something went wrong'
Apply replacement to example:
'in about 1790-1795 something went wrong'
```
- Buttons**: "Ok" and "Cancel" buttons at the bottom.

Rewriting... – this heading reminds you, which “fixer” you are working on.

Sample text – this is used as the input if you click the Test / results button.

Recognizer pattern – the regular expression used to check input text for a match and optionally to extract parts of it for re-use in the replacement pattern.

Replacement pattern – the expression used to replace the input text when the recognizer pattern finds a match. NB: the entire input is replaced; so, if you want to keep part of the input, the patterns need to be designed with that in mind.

Test / results button – click this to have Ged2Reg attempt to build a Regular Expression instance from the text in the recognizer box, run the input through it, and apply the replacement pattern. The results, step-by-step, are listed in the (read-only) text box. Note that there are many things that can go wrong here: the syntax is quite exacting. If, for example, the text in recognizer field is not a valid regular expression, an error message is all that will be shown.

Ok button – the current values here for the recognizer and replacer are applied to the current settings for the appropriate fixer, depending on which button you used to access this dialog; and this dialog is closed.

Cancel button – any values entered or changed here are discarded, and this dialog is closed.

Explanation of the Example

The fixer shown in the screen shot was developed for use with a GEDCOM file in which many event descriptions had, somehow, suffered a particular kind of “damage” in which a required blank space between a date (year) and the following word had been dropped. The overall “logic” of the fixer is to find occurrences of that problem, breaking the description into two parts in the process; and then put them back together with the space between them.

Here’s a short explanation of what the parts of the regular expression do:⁸

- **(?i)** - this sets a *mode* of case-insensitive operation; so that capital letters and lower-case letters are considered to be the same
- **(?<part1>^.*?\d{4})** - this defines a *named capture group* that recognizes and picks out the left “half” of an occurrence of the problem situation
 - **(?<part1>)** - the outer parentheses indicate that everything inside goes together as one *capture group*; then **?<part1>** assigns the name part1 to the group
 - **^** - this is an *anchor* that indicates that the pattern starts checking for a match (and collecting the characters in the group) from the very beginning of the text being considered
 - **\d{4}** - this is a *character class* with a *quantifier*, together indicating that there must be four numeric digits in succession to make a match
 - **. *?** - this indicates that, after the beginning, we accept anything (.) any number of times (*) up to but not including (?) an occurrence of a four-digit sequence
- **(?<part2>[a-z]{2,}.*\$)** - this recognizes and picks out the right “half” of an occurrence
 - **(?<part2>)** - the outer parentheses indicate that everything inside goes together as one “group”; then **?<part2>** assigns the name part2 to the group
 - **[a-z]{2,}** - this is a *custom character class* (**[a-z]**) that matches any letters, with a *quantifier* (**{2,}**) that indicates at least two in succession are required to make a match
 - **\$** - this is another *anchor*, one that matches the very end of the input text.

⁸ The terms in italics are parts of the technical vocabulary of regular expressions; we are not going to define them here, but you may find more about them when reading up on the subject in other resources.

- `. *` - this is a *character class* (dot, any character) with a *quantifier* (`*`, any number of times) indicating that here we accept anything from the previous matching sequence up to what follows (i.e., up to the end).

Because there is no space (nor anything else) between the two “parts”, the matching input must be continuous. So, “something 2021and then something else” would match, but “something 2021 and then something else” would not.

Finally, we have the replacement pattern:

- `${part1} ${part2}` – the logic in Ged2Reg that uses these patterns ensures that the replacement is applied if and only if the whole thing matches. This particular replace pattern is used to put the two “halves” back together, in the original order, but with one space inserted between them.

This is just one, relatively straightforward example of solving a specific problem. If we only wanted to keep one half of the matching input, we would simply have left the other half out of the replacement pattern. The potential variations are infinite.

You may have realized that you could also (instead) solve this problem in the word processing application. In many cases that may be faster and easier! The choice may depend on factors including your comfort level with this technology,⁹ how many cases need fixing, how long the report(s) are, and how many different reports you plan to create from the same input.

⁹ No, “sorry”, we are not in a position to “help” write the regexes you may need.

Supplemental Information

Styles

Ged2Reg makes use of “styles” to control the formatting of characters and paragraphs in the output. This is an essential feature for enabling effective use of word processing programs to extend and customize the output. The contrast to the alternative – direct application of format properties (such as font, font size, font weight, indentation, and on and on) to separate pieces of text – is stark. With styles, you may decide quite late that perhaps it would be better, say, to use a different font for the children... and change that throughout the document in one simple step; without styles, you may be stuck re-editing tens or hundreds of places one-by-one.

Most of the styles in the Ged2Reg output – both names and properties – are based on recommendations from the NEHGS, recommendations that are embodied in a sample document they make available on the Internet.¹⁰ And, to facilitate editing, additional recommended styles from that resource are included in the document even though they are not used in the content generated by Ged2Reg.

There are some divergences from the NEHGS style recommendations in Ged2Reg output. First, the suggestions indicate that each main person paragraph, except for the main person name, is to be in *Normal* style. We consider this to be an error in judgement and decline to comply. *Normal* style in Word is the basis of all other styles. Putting part of the document in *Normal* style makes it impossible to restyle those elements without changing everything else as a side effect. *Normal* style should, we firmly hold, never be used directly on any content in a “substantial” Word document. Secondly, the NEHGS recommendation suggests that generation numbers should be made superscript by directly applying the property to each one where it appears. Again, we consider this to be an error in judgement; Ged2Reg uses a style to control this property, making it accessible to simple, global change as it should be.

There are a very few places where Ged2Reg (in docx output files) may control the appearance “directly” by applying properties to text rather than using a style. That includes “divider lines” (which are actually “borders”) and the optional style of the notes introduction line. In odt files, so far as we can see, there is no mechanism for direct application of properties to text, so Ged2Reg generates additional styles to effect the desired appearance in those cases.

The principal styles used in the Ged2Reg output files are:

- **MainPersonText** – paragraph style for the main person paragraph; introduced by Ged2Reg and used instead of *Normal*.
- **MainPersonName** – character style for, of course, the name of the person at the start of a main person paragraph.
- **BodyTextIndent** – paragraph style that is specified in *MainPersonText* as the *Style for following paragraph*. Thus, if you want to extend a main person’s story, just position the cursor at the end of that person’s main paragraph and press the enter key; the new paragraph will be in the (this) recommended style, ready for you to type.

¹⁰ See [Submission Guidelines](https://media.americanancestors.org/uploadedfiles/media/the_register/register-template_rev2017-aug.doc) page, which mentions the “template”; direct link to the docx file: https://media.americanancestors.org/uploadedfiles/media/the_register/register-template_rev2017-aug.doc (links verified as of 24 March 2021).

- **BodyTextNotes** – paragraph style, based on BodyTextIndent. Testing with GEDCOM files that include extensive notes suggested that they may look better on the page if the properties of the text are adjusted, for instance to use a smaller font. This style was introduced to apply that observation, while keeping it a simple process to apply a different choice within a given document.
- **GenerationNumber** – character style introduced by Ged2Reg, used for the *generation number* superscripts.
- **KidsIntro** – paragraph style for the text between the main person and the children; “Children of...”.
- **Kids** – paragraph style for the children paragraphs. This style includes the all-important tab stops used to align the child numbering. If your numbers become large and you want to allow more space, adjusting these tab stops in the style is where you would do that. Note that this can be extremely tricky in OLO; be sure to save your work first!
- **ChildName** – character style for the child’s name as it first appears in a child paragraph.

Marriages vs. Names

The naming convention for a married woman in the Register format is pretty clear: they expect it to look like this: GivenName (SurnameAtBirth) (PreviousMarriedSurname)... MarriedSurname. But, so far as we understand it, GEDCOM does not have an unambiguous way to convey the answers to some basic, and quite relevant, questions about a family unit: were the couple formally married?; and, what surname did the woman use when married?

Then, when we reach into the Medieval period, the matter is further complicated by the fact that the binomial convention had not yet fully developed. Technically, this aspect is better handled in GEDCOM. Within the GEDCOM file each person’s surname is set off by slashes within the content of a name record, like this:

```
1 NAME John /Smith/
```

and when a person had no surname there should be nothing between the slashes.¹¹ In practice, though, this may still be a problem area. For instance, we may (should) know, for example, that in the name “Alice of Northumberland”, the final word is *really not* a surname “Northumberland”; this woman did not have a surname in the modern sense. But if you type that name into some genealogy programs, they will take “Alice of” to be the given names, and take the rest for a surname; and they will export it into GEDCOM that way. You may be able to fix this, if you care to. For example in FTM you would need to make a second edit of the name by clicking on the (hidden) button at the right side of the input field for Name. That opens a popup where you can cut *Northumberland* out of the surname sub-field, (where FTM had helpfully split it out), and paste it into the end of the given name subfield, leaving the surname part empty. If you do that, the GEDCOM export from FTM is correct. Got it? Been doing that?

This all matters here because there are several places where Ged2Reg needs to try to assemble a “conforming married name” from the information it finds in the GEDCOM. To do that, Ged2Reg

¹¹ There are also separate tags for given names (GIVN) and surname (SURN) that can be optionally included under the NAME tag in a GEDCOM. This, too, would seem to be unambiguous; however, use of these optional tags does not appear to be common.

assumes that all couples were married and that the wife always took the husband's surname, if he had one. These assumptions are not normative. They do not represent a developer's belief about the way things "ought to be done". They are not a political or religious or moral or cultural opinion. They are simply the choices leading to data manipulation logic that is going to give the correct results for the greatest number of actual cases in the ancestry of American researchers in the time span from, say, about the mid-first millennium to at least the mid-20th-century.

The user is advised to know their data, and to be aware of the potential need to edit the output of Ged2Reg to adjust the text for any cases where this convention may result in a "name" that does not properly represent the person. For example, you may need to edit in cases where a couple were not formally married, in cases where the woman retained her birth surname after marriage, in cases where one or both of the couple took a new compound surname after marriage, and so on.

Further, if your report reaches into the Medieval period you may need to edit to deal with improperly extended "surnames". When Ged2Reg detects that a person's surname subfield is affirmatively empty, it assumes that the person(s) involved did not have binomial personal names and does not try to apply this logic to forming a married name. But if you let e.g. FTM put e.g. *Northumberland*, or *Conqueror*, and so on, in the surname sub-field, that is what Ged2Reg is going to see. Finding a name like "Maud (Flanders) Conqueror" in your output may be amusing, but it would not impress any editor or scholar; nor would this misconstruction be Ged2Reg's problem.

Regular Expressions

Regular expressions can be thought of as a "grown up" elaboration of the familiar wildcard search using "*" for anything, as in "census*.jpg" to find all JPEG image files that start with the word "census".

A particular regular expression can be simple, or so complex the person who wrote it can't understand what it does 10 minutes later, or anywhere in between. Explaining how they work is "beyond the scope of this document".

Here is a link to Microsoft documentation: <https://docs.microsoft.com/en-us/dotnet/standard/base-types/regular-expressions>. It is part of a site intended for programmers, but this particular topic is covered in a way that might be helpful to non-programmers. Googling for 'regular expression tutorial' may also help. And, if you get serious about the topic, the desktop application "Regex Buddy" (license fee required) is a tool that "we" have used to build and test Regular Expressions; see <https://www.regexbuddy.com/>.

There is also an experimental online tool that may¹² be able to build a regular expression for you if you can supply multiple examples of what you are trying to match: <http://regex.inginf.units.it/>.

Miscellaneous Topics

Source titles. According to conventions, some source titles should be in italics (published books), some should be in quotes (published articles), and others should be neither. Ged2Reg does not have the information that would be needed for it to conform to these conventions; so, it does not. If the input has titles in quotes, they should appear in the document; if the input has titles wrapped in "tags"

¹² We did not succeed at this; but, we did not try very hard, because we know how to write regular expressions, to a "medium" level of complexity.

indicating italics (<i> ... </i>) then the text so marked will be in italics (in odt files) or plain text with the tags removed (docx).

Test out your editorial process. Editing and assembling documents can be a complicated process. Choosing some Ged2Reg options may make this easier, or more difficult. For example, bracketed footnote numbers look great; but, manually inserted footnotes will not be automatically bracketed to match. For another, endnotes can be appealing in getting the documentation “out of the way” and more of the basic content on each page; but, they can be extremely hard to work with if the Ged2Reg output is only part of the finished work. And (lots of potential issues with footnotes) footnote references may be output to pdf as links that let the user navigate within the document; but, not all variations work equally well and the two families of word processing programs (Word vs. OLO) do not work exactly the same.

The more elaborate your intended project is, the more important it is to test your entire process before investing a great deal of time and effort working on particular artifacts. If your project involves just one Ged2Reg report that you expect to edit a bit and then export to a pdf file, you *probably should test* out the “export to pdf” with the Ged2Reg report before you do a lot of editing work. If your project involves assembling an extended document or book from one or more Ged2Reg reports, others chapters written separately, indexing, table of contents, etc. etc., you *certainly should test* your anticipated assembly and final output process using drafts that you invest as little work into as possible, in order to make sure your process will work and that you know how to get the tools you are using to do what you want.

Places where Ged2Reg can’t (or, just doesn’t) conform to conventions. If all of the children in a group were born (or baptized) at the same place, that detail is supposed to be stated on the “intro” line and then omitted from each child. Ged2Reg does not do this. Also, the convention suggests abbreviating month names and state names, if five characters or longer, in child paragraphs. Ged2Reg does not do this.

Multiple parent relationships. If a person in the GEDCOM file has multiple parent relationships, such as when a woman’s children from her first marriage were adopted by her second husband (and, you created those relationship links in your genealogy application), then the person may appear in the report in an unexpected place, or in more than one place. You should review the output for such cases and edit as necessary.

URLs (web links) in plain text. If your source citations include web links in “unexpected” places (i.e., not where FTM or RM7 puts them, but instead in a field like “page”), they should still be recognized and formatted according to the option selected; but, they will not be wrapped in the additional text that is included by default in the *Citation Options*.

Potential difficulties in editing. If you use “See Note..” to reduce the repetition in footnotes, and then do certain kinds of editing in the document using your word processing program, some difficulty may arise. You can’t tell when looking at a footnote, whether or not a later one refers to it. So, it is possible that you could delete a note that is referred to later, or change it to specify a different source that may not apply to the later points in the document, without noticing the impact at that moment. So: “be careful”.

Large documents. If your report extends to thousands of persons, tens of thousands of citations, tens of thousands of index entries, you may have some difficulty working with it in the word processing

application. Testing suggests that this is especially true of the free applications, Open Office / Libre Office (OLO). As noted elsewhere, we find that OLO simply cannot open docx files with several tens of thousands of footnotes.¹³ We also “sometimes” have difficulty with later steps, e.g. updating the index (may seem to hang; try waiting it out, or try again without “married names” indexing option?) and with export to pdf (try waiting it out; we have seen this step take an hour or so...).

Working with Libre Office. We respect and appreciate the efforts of the volunteers who work on this free software, and understand why many users prefer it to the fee-based alternative. But, we also note that it is not uncommon to have difficulty using Libre Office (and/or Open Office), especially on large documents comprising thousands of pages and tens of thousands of footnotes and index entries.

The biggest problem seems to be a certain fragility in the “pagination” process (where the program works through the text from start to end and lays the content out on pages; which is greatly complicated by footnotes). In our experience it is critically important, on large files especially, to give Libre Office plenty of time to “settle down” after opening a large file before doing anything.

In a recent test of two alternative courses of action starting from the point of opening a large (ultimately, about 2700-page) odt file using LibreOffice 7, we noted that LibreOffice Writer displayed the first page and enabled the UI in a fairly short time, about 15 seconds. At that point the status bar at the bottom of the window said the total number of pages was 701.

If at that point we “jumped” to the end of the file using the ctrl-end key combination, Writer took a little while to “complete” the pagination process, but it did not do it correctly. Instead, when we scrolled through the file we could see that beginning shortly after page 701, Writer had started page-breaking arbitrarily, leaving many pages half-full or less, and ultimately doubled the total page count to about 6000.

If instead (closing the program and starting over) we waited another minute or so after Writer first displayed that first page and the low page count, we could see that Writer resumed its pagination process. The count displayed in the status bar started rapidly spinning up from 701, coming to rest again in the 790s in another minute or so. If at that point we jumped to the end of the document and then waited again, in a couple of minutes Writer apparently finished pagination and display a page count that we think is correct, around 2700. And, scrolling through the document shows that the half-empty page problem had not occurred this time.

If your file has already been paginated incorrectly the situation may not be hopeless. The supposedly correct way to get LibreOffice to repaginate the whole file is to select from the menu, Tools > Update > Update All; but, in our experience, when this action is performed on a large file Writer just hangs. But if instead we scroll to about that critical point (page 701+, in this case) and look for the first occurrence of a half-empty page, we could fix it. The trick was to position the cursor at the very end of the body text (meaning, not footnotes) on the half-filled page, hit the delete key once, and hit the enter key once. Writer then positioned the following text correctly so as to fill that previously half-empty page, and continued to repaginate to the end, getting rid of the half-empty pages... provided we gave it time to finish.

Finally, it seems to be important to turn OFF the *Track Changes* option on the *Edit* menu, especially when trying to update the index in a large file. And, we suspect that Writer’s *Save Autorecovery*

¹³ That is one of the key reasons that Ged2Reg was extended to provide the option of outputting odt files.

Information option causes problems when the automatic save happens in the middle of a long-running operation (such as, update index) on a large file. It may be preferable to turn it off: starting with the menu bar, you reach the option setting by selecting *Tools > Options > Load/Save > General* .

For all that, we have not succeeded trying to *Update Index* in LibreOffice, on a file with 35,000 index entries. The latest version of Open Office was able to build that index (at least once, though not every time we tried it; *sigh*); but we did not succeed with everything else we tried in that version of the software.