



Search of Persons with defined Source data

Util - SourPerson - Diedrich Hesmer, Herrenberg

Create List of Persons for defined Sources

Base data + Source selection | Selection of Persons | Output definition | Explanations

Program scans a ged-file for persons with defined source data and other selection criteria and creates a list of persons found including defined stored data. Alternatively the search may include families.

Scan ged-file
C:\Ofb\1\ages\schreiner\schreiner-hiemer_u.ged

Selection of sources

SOUR Scan: INDI, INDI + FAM

SOUR Type: Record, Embedded

SOUR Levels: only level 1, all levels

Selection of:
- SOUR Scan: Record type for source checking
- SOUR Type: Type of storing sources
- SOUR Levels: the levels to be checked

Search tags + texts - multiple separated by semicolon, Wildcard chars "*" # "?" allowed

Selection of search tags and entering of search texts. Select tags from top to bottom. The "Search tags" will be compared by "and", the "Search texts" separated by semicolon will be compared by "or". The "Search texts" may be in small letters, the original will be converted accordingly.

PUBL	*.org*
TITL	*agr*
REPO.NOTE	zwilling*
REPO.NAME	*ance*;genea*

Selection file - *.ussd
C:\Ofb\1\ages\schreiner\sp-test1.ussd

New Info Start Edit Log Start Browser Close

© Diedrich Hesmer, 2017-2017 Version: 0.4.4

File saved

Functionality

This program scans a ged-file for persons with defined sources data and other criteria for the person selection and creates a html file of the persons found with the desired data. In a run, the sources of INDI and, optionally, FAM records can be scanned. Persons with at least 1 SOUR structure (data record or embedded data) matching all defined SOUR search criteria and which additionally meet all defined search criteria for the person selection will be listed. Records that do not contain all defined tags and criteria are ignored.

Fields with missing or erroneous entries are marked red, which allows so far no analysis.

A log file is created during processing, which can be opened with any text editor.

Description

The description is divided into the program sequence, the search criteria and output definitions, the buttons and their functions and the log-file and a report example is shown.

Program sequence

After starting, the program above screen will show up.



GSP Utility – SourPerson

There are 4 tabs for definition of the selections and output of the findings and for "Explanations", which should be read and understood. They include more details as described here.

The line at the bottom is a message line.

- At the first call, all fields are empty resp. set to the system values.
- The ged-file and the selection criteria and options must be defined.
- Save all the definitions. By clicking the diskette symbol the standard dialog starts. Enter an easy to remember name. The data will be stored in an ussd-file.
- By "Start" the processing starts. The search criteria will be compared with the data of the ged-file. In case of compliance, the person will be counted as finding.
- The persons found will be stored in a html-file, to be opened by "Start Browser".
- During a next start of the program it will search for the last used ussd-file. If found, the settings of the last "Close" will be used and the fields filled accordingly. They may be modified before pushing "Start".

Search criteria

These defines the scan for the sources and persons. You have to select/define:

Source selection

- Select records to be scanned - INDI or INDI + FAM.
- Define if the sources are stored within the ged-file as records or embedded text.
- Define if for sources only level-1 or all levels should be scanned.
- Select at least 1 of the 6 (for records) resp. 3 (for embedded) fields for search tags resp. tag combinations:
 - for records:
 - TITL - from SOUR record
 - TEXT - from SOUR record + SOUR sub-tag DATA.TEXT
 - NOTE - from SOUR record + SOUR sub-tag NOTE
 - PUBL - from SOUR record
 - AUTH - from SOUR record
 - DATA.AGNC - from SOUR record
 - DATA.NOTE - from SOUR record
 - REPO.ADDR - from REPO record
 - REPO.NAME - from REPO record
 - REPO.NOTE - from REPO record + REPO sub-tag NOTE
 - for embedded:
 - TITL - from SOUR xxx
 - TEXT - from SOUR sub-tag TEXT
 - NOTE - from SOUR sub-tag NOTE

The fields must be filled from top to bottom. After an empty field, no more filled one may follow.

- For every selected field enter the search text (lower case is ok). Multiple search texts to be separated by semicolon (will be compared by "or"). The wildcards * # ? are allowed. If only 1 asterisk is entered, all data will be used. If, however, no asterisk is specified, only the data that exactly corresponds to the specified text is used. For scanning all texts will be converted to lower case..



GSP Utility – SourPerson

Selection of Persons

The setup of the screens for selection of Persons and Output definition are shown below on the right exemplary.

- Enter search text for the data fields to be scanned. The fields may stay empty and therefore won't be scanned. Here also for e.g. notes a single * may be entered. This means that all persons who have a NOTE tag in the data record, independent of the content, are transferred to the data to be checked. For NOTE optional the search may be limited to only level-1.
- For selection of places and dates the events birth/christening, death/burial and marriage exists. Only for missing data for birth resp. death the christening resp. burial data will be used.
- For selection of places, the text of tag PLAC will be used. Optional the text up to the 1st comma only will be used. Here also the place of "Citizenship" can be selected. This requires the selection of the storage method 1..5 used in the ged file and the text used (tag name or EVEN/TYPE text) must be entered. The system offers texts for each method that can be overwritten. The scheme is displayed below by selecting the method.
- For the search texts of the data fields and selection of places, the same explanations as for selection of sources applies.
- For selection of dates 2 fields (from + to) exists. They must be entered in the format "YYYYMMDD" (year month day). If 1 value is entered, the comparison will be done by "=" (exact). This value also may be "19700000" for the year 1970 or "19700700" for July 1970. Internally a 2nd value "19701231" resp. "19700731" as end value will be created.
- If 2 values are entered or internally created, the comparison will be done by ">= value1" and "<= value2". For the comparison only the date value of the DATE text will be used. All other text and the uncertainties e.g. AFT, BEF, ... will be suppressed. In case of 2 values (date ranges) only the 1st value (FROM, BET) will be used. Beside the Gedcom format also German date formats "17.8.2012" will be processed.
Note: For date comparison the DATE text of the ged-file will be calculated as follows:
DATE 1970 -> 19700000, DATE JAN 1970 -> 19700100, DATE 10 JAN 1970 -> 19700110

A finding is given if a match with ALL defined criteria was found in the ged file. The details of the logic you'll find in tab "Explanations" > "Logic of Selection".

Output definition

Rules for the Output definition:

- Selection resp. entry of text of the offered options.
- Within the output after the NAME with SEX + RELI following groups will start in a new line: the groups BIRT with CHR, DEAT with BUR1, SOUR and the group of the remaining level-1 tags, where NOTE is placed at the beginning of this group. Below of the persons data optional the FAMS data will follow, first MARR with DIV, then as described above.
- The output may be sorted and the font size reduced. Level-1 tags will be listed in bold, italic and gray, Tags of level-2 and higher in italic and gray. The tags SOUR, NOTE and FAM are highlighted in color.



GSP Utility – SourPerson

- For sorting any special characters may be deleted at Surnames and Given names.
- For handwritten entries, blank lines can be inserted for each listed person at the end of the output.

Rules for Definition of Output data:

- Decide if tags for output should be defined, otherwise all tags except the following will be listed: GIVN MAP SUBM SURN FONE ROMN CHAN _CREAT _LOC _UID.
- If they should be defined, decide if a group of tags should be suppressed for output or if additional tags should be added to a base group.
 - Always suppressed will be the tags listed below. Further tags may be entered. GIVN SURN AFN ANCI DESI MAP OBJE RFN RIN SUBM CHAN _CREAT _LOC _UID FONE ROMN HUSB WIFE CHIL
 - As base tags always will be listed.
 - for level-1 the tags NAME BIRT CHR DEAT BURI MARR DIV SOUR SEX RELI FAMS and optional NOTE.
 - for level-2 the tags DATE, PLAC, SOUR, TYPE and optional NOTE.

Further tags may be entered, separated for level-1 and level-2.

Note: Tags suppressed as default may be forced for output by this feature.

- Multiple tags to be separated by semicolon, no space allowed.
- By the option "include FAM data" for all FAMS Ids the important FAM tags will be listed, otherwise only the FAMS Ids where SOUR data are found.

The output list "xxx_soep.html" ("xxx" = ged filename) will be stored in the folder of the ged-file and may be viewed by any Browser.

Logic of Selection

The checking will be done in following sequence:

- Step 1: Those persons are searched for, where the defined search texts of the source selection were found.
- Step 2: From these only those will be listed, where all defined search criteria for the selection of persons applies.

Rules for selection:

- All selected search tags and data fields are evaluated separately.
- Search text for selection of sources and persons separated by semicolon will be split into individual pieces of text.
- Every individual search text will be compared to the according text of the ged-file.
Note: CONC and CONT lines will already be concatenated to their base tag during data reading of the ged-file
 - CONC directly, CONT by ", "
- For the first match found, the result for the search text resp. data field is considered "true".
- For the 3 date and 4 place checks the result is considered "true", if already 1 of the selected checks are corresponding.
- The test results obtained for each search text resp. data field are now compared with "and".
- If this result is "true", the person is selected - optional also the families data linked by the FAMS Id."

Horizontal is compared with "or". In case of multiple text pieces (search texts separated by ;) and options for selection of places + dates only 1 comparison must be true.

Vertical is compared with "and" and every selected element must be true.

If a search tag or data field, defined for searching, is not included within a record, the comparison will be set to "False". Therefore this person is NOT included in the output list. even if all other criteria comply.



Buttons

- [Symbol with an open folder] Loads a ged-file.
- [Edit] Pushing by the left mouse button opens the ged-file by the text editor for viewing. Pushing the right mouse button starts "GedShow" with the ged-file.
- [x] Deletes all entries of the search criteria.
- [?] Opens the description.
- [New] Clears all definitions.
- [Info] Shows the folder where the ini-file is stored (contains all settings at the last "Close").
- [Start] Starts the processing.
- [Start Browser] Opens the html-file by the Browser.
- [Edit Log] Opens the log-file by the text editor.
- [Close] Writes the ini-file with all settings and closes the program. Closing the program by the little "x" at the upper right corner of the screen will not write the ini-file.)

Extract of a Log-file

The file lists:

- Number of lines of the ged-file.
- Number of records found per type.
- Number of records selected per type.
- Processing explanations.
- Output information.

```
Created at 08.07.2017 16:11:43 - SourPerson version: 0.4.1
Base: "C:\Ofb\1\ftm\test\Testfamilie.GED" from 07.07.2017 19:15:25
08.07.2017 16:11:43 Start Scan
Lines read: 316
Records found: 17 INDI, 6 FAM, 7 NOTE, 2 SOUR, 1 REPO
Records selected: 2 INDI, 1 FAM, 3 NOTE, 1 SOUR, 0 REPO
Record Ids replaced by text
08.07.2017 16:11:43 Reading ged-file for output
2 INDI read for output
List with 2 entries in "c:\ofb\1\ftm\test\testfamilie_sope.html"
08.07.2017 16:11:43 Analysis completed
```



Report example

With inserted empty lines for notes.

Testlauf

```
# INDI
1 X1234567890 NAME: Proband Vorname /Mustermann/, SOUR: TITL: Standesamt Musterhausen, REPO: NAME: TTTTT, NOTE: Standesamt Musterhausen Urkunde 4711; SEX: M;
REL: P: Evangelisch, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE: Karl Autor, Testring 1, 12345 Musterhausen, Karl@@Autor.de, REPO: NOTE:
Familien in Musterhausen
BIRT: D: 31 DEC 1930, P: Musterhausen, SOUR: TITL: Standesamt Musterhausen, REPO: NAME: TTTTT, NOTE: Standesamt Musterhausen Urkunde 4711
DEAT: D: 15 MAR 2016, P: Musterhausen, SOUR: TITL: Standesamt Musterhausen, REPO: NAME: TTTTT, NOTE: Standesamt Musterhausen Urkunde 999/2016, CAUS:
Herzinfarkt, BURF: D: 18 MAR 2016, P: Musterhausen, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE: Karl Autor, Testring 1, 12345 Musterhausen,
Karl@@Autor.de, REPO: NOTE: Familien in Musterhausen; NOTE: Diese Notiz ist nur ein Test-Beispiel
ALIA: Proband d. I.J.; TITL: Dipl.-Ing.; OCCU: P: Sohrreiner, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE: Karl Autor, Testring 1, 12345 Musterhausen,
Karl@@Autor.de, REPO: NOTE: Familien in Musterhausen; BIAPM: D: 2 JAN 1931, P: Musterhausen, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE:
Karl Autor, Testring 1, 12345 Musterhausen, Karl@@Autor.de, REPO: NOTE: Familien in Musterhausen; REFN: 1; ADDR: ADR1: Teststr. 55, CITY: Musterhausen, POST:
12345, CTRY: Deutschland; FAMS: @F3@
FAM: @F3@; MAARR: D: 31 MAY 1960, P: Musterhausen, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE: Karl Autor, Testring 1, 12345 Musterhausen,
Karl@@Autor.de, REPO: NOTE: Familien in Musterhausen Nr. 99, SOUR: TITL: Standesamt Musterhausen, REPO: NAME: TTTTT, NOTE: Standesamt Musterhausen
Urkunde 333/1960; NOTE: Ehefrau hat als Witwe nochmals geheiratet

2 104 NAME: Ehefrau /Musterfrau/, SEX: F
REFN: 4; FAMS: @F3@
FAM: @F3@; MAARR: D: 31 MAY 1960, P: Musterhausen, SOUR: TITL: Familien in Musterhausen, AUTH: Karl Autor, NOTE: Karl Autor, Testring 1, 12345 Musterhausen,
Karl@@Autor.de, REPO: NOTE: Familien in Musterhausen Nr. 99, SOUR: TITL: Standesamt Musterhausen, REPO: NAME: TTTTT, NOTE: Standesamt Musterhausen
Urkunde 333/1960; NOTE: Ehefrau hat als Witwe nochmals geheiratet
```

Author: **Diedrich Hesmer, Herrenberg**

Created on 08.07.2017 with [SourPerson](#) @ V: 0.4.2 from Diedrich Hesmer
based on data from "testfamilie.ged"

Selection criteria:

Base data + Source selection
SOUR Scan: INDI + FAM
SOUR Type: Record
SOUR Levels: all levels
Search tags + texts:
NOTE - "musterhausen"

Selection of Persons
Selection of Persons data
Surname: Muster*

Output definition
Output options
include FAM data
Define tags for output: suppress
Suppress tags in output: FAMC,EVEN

Versions

(N = New, C = corrected, M = modified)

- 1.0.0 – 02.07.2017 – N: Initial publication
- 1.0.1 – 19.03.2018 – N: GedShow included
- 1.0.3 – 08.05.2018 – N: "?" Show Docu