Importing and Exporting

GenScriber will import and export several different formats. Some formats are handled by GenScriber, other formats require the UNO app found in Libre Office and Open Office.

Imported files are always converted to csv files. These are utf-8 by default. If a different character set is chosen in "Preferences" the import export will attempt to convert from the selected code page.
The exception to this is Gedcom. Exported gedcom files are always UTF-8.

When importing gedcoms, the file header is used to determine the character set used in the file.
GenScriber will import Ansel, UTF-8,   ASCII   and ANSI .

ANSI is not actually allowed in gedcoms, but it is sometimes used and when it is used, it needs to be handled differently.
If the character set in preferences is utf-8, then the windows-1251 codepage will be used.
If a character set other that utf-8 has been chosen in preferences, then that codepage will be used instead.

Merging files

File Merge.

Multiple files can be opened and merged automatically. You can Import and Merge at the same time.
File types are allowed to be different, but all merged files must use the same template layout (identical column titles in the same order).

How to select multiple files to open/merge.

When in the '   **Open File**   ' dialog, hold down the '   **Ctrl'**   or   **'Shift'**   key to select multiple files.
Files will be opened and concatenated in the order found on disk.

The name of the first file opened will be used as the filename, with '-new' appended.

The resulting file will be CSV format.

example file1.xls +file2.ods +file3.csv = file1-new.csv

Setting up uno import export

Setting up import/export using unoconv.

GenScriber 2.1.5 and above will import and export several extra file formats, including ods and xls.

The 'Export with uno' option must be selected in '   **Tools>Preferences>Export**   '.

Important: For this to work, you must have either LibreOffice or OpenOffice installed on your system.
On Linux systems this is not a problem, either LibreOffice or OpenOffice is usually installed by default, and it should work out of the box.

Windows users will need to install LibreOffice (or OpenOffice) manually. If you don't already have it, you can download it here…
  http://www.libreoffice.org

Below is a check list to help make the export functions work for you.
1. Install LibreOffice or OpenOffice (version 3 or above) .
2. Start GenScriber.
3. Go to Tools>Preferences>Export'.
4. Check the 'Export with uno' option is selected.
5. Check that there is a path entry to python.exe.
6. If no path, click on 'search' and wait until the path appears.
6. Close Preferences.
If all of the above is ok, you will be able to open and export several different
spreadsheet formats.

If the import/export functions don't work for you.

Windows 7 and 8 users.

Win 7/8 blocks almost any file it does not own. That is just about every file you introduce into your system. Unless you have already turned this off, there is a good chance that the genscriber support files will be blocked.
If things don't work as expected, this is probably your problem.

Windows needs to know where to find python.exe. This will usually be something like
"C:/Program Files/LibreOffice 4/program/python.exe".
When **'Preferences'** is opened, genscriber will attempt to find the correct path and add it for you. The path will be displayed in **'Tools>Preferences>Export>Path to python'** . If you see a path here you probably don't need to change it.

If the path is empty, then you need to tell GenScriber to search for python.exe.
When in **Tools>Preferences>Export** click on the ' **search** ' button (*to the right of the path box*). GenScriber will carry out a deep search and add the path for you.

The GHF file format

The GHF Header
The GHF header is a simple and reliable way to add extra information to csv files. It is placed at the top of a normal CSV file, above the title row.
Unlike some info headers, ghf is compatible with spreadsheets.

The GHF header uses open and close tags **<INFO>** and **</INFO>**
This allows parsing applications to easily extract the information.

The <INFO> tag starts the header, and is placed in the first column of the first row.

The first column of rows following the <INFO> tag may contain the information items.
Info items can contain anything except commas and quotes.
A maximum of 10 info items(rows) may be place between the tags.
Tags and info items are **always in the first column** . Other columns are ignored.

The header is then closed with the </INFO> tag. This must also be in the first column.

Column Titles must be on the row immediately after the closing tag.

Example header:
   <INFO> ,,,,,,,,,,,
Church: St. josephs,,,,,,,,,,,
Parish: Charlestown,,,,,,,,,,, Source: Parish records (original),,,,,,,,,,,
Transcriber: Les Hardy,,,,,,,,,,,
Email: les@webmayo.com,,,,,,,,,,,
   </INFO> ,,,,,,,,,,,
Surname,FirstName,Day,Month,Year,Abode,County,Aged,Rela,Status,Extra,Occupation

Note: the extra commas match the title row, and are there to ensure compatibilty with spreadsheet applications.
Adding The GHF Info Header
The header information you enter into the header info fields will be added automatically to the worksheet when exported as GHF.Use the '.ghf ' file suffix for GHF files.

The Information
Ten fields are available.
Any information you enter here refers to the current worksheet, and if the worksheet is exported as .ghf, it will be saved in the file header.

When a new file is opened, all previous header info fields will be cleared.
If a .ghf file is opened, any header info it contains will be shown in the header information.

If you save as .csv, the header information will NOT be saved into the csv file, but the information fields will remain until the file is closed.

Field Titles
You can add/edit the field titles by simply placing your mouse cursor in the grey area above a field, and left-click once, then type in the field title.
Field titles are retained when you quit genscriber, and will remain until you change them.

FreeCEN format

Worksheets can be saved with the FreeCEN file format'.
If this menu option is unavailable, go to ' **Tools/Preferences** ' and under the ' **Export** ' tab select ' **Export FreeCEN** '.

The Header Information
The FreeCEN header information will be included in the header of files exported as 'FreeCEN'
The FreeCEN file is saved with a ' **.cen** ' suffix.

You will need to change the file suffix to '.csv' before sending it to your coordinator.

Census Piece: This is the census piece being transcribed. example: RG121780
Registration District Number: The registration district number you are working on.
Transcriber Code: Enter your transcriber code here.
Transcriber Name: Enter your name here. (optional)

Each time a FreeCEN (.cen) file is opened, the header information is retrieved (if it exists).
When a FreeCEN (.cen) file is exported, the current header information is saved in the file.
It is quite acceptable to open a csv file, edit the FreeCEN header info, and then export as FreeCEN.

FreeBMD Format

Worksheets can be opened and saved with the FreeBMD 'flat file header'.

The following explanation was created from information found on the FreeBMD website…
The "Flat" file format is designed to cater for datasets stored in a database or spreadsheet with all data for each entry on a single line (hence the name "flat").

The essence of the flat format is this, so long as all required data is available, it can be submitted in almost any layout. A line at the top of the file defines what goes where.

Example of Flat File Format
+INFO,jverge@assendon.demon.co.uk,pw,ONENAME,BIRTHS ,UTF-8
+FORMAT,Surname,Firstname,District,Volume,Page,Quarter,Year,Type
VERGE,John,Ringwood,VIII,205,Mar,1838,B
VERGE,George,Ringwood,VIII,198,Jun,1841,B
VERGE,Mary,Fordingbridge,VIII,181,Mar,1843,B
VERGE,Joseph,London,II,177,Sep,1843,B

Column types
Each column type may have alternate names, which are shown in brackets after the standard name.
Surname The surname of the subject of the entry.
FirstNames (FirstName,Forename) The first names of the subject of the entry.
OtherName optional. If present, this is appended to the firstnames field.
District The district in which the event was registered.
Volume The volume number in which the event appears.
Page The page number of the event.
Quarter The quarter of the entry (i.e. Mar, Jun, Sep). You can use QuarterNumber instead.
QuarterNumber is the number of the quarter of the entry (in the range 1-4, 1 is Mar and so on).
Year The year of the entry.
Type (Event) optional The event type - one of: BIRTH, MARRIAGE, DEATH.
N.B. Event types cannot be mixed in a single file. This field is therefore only useful as a crosscheck.

Additional Optional Columns
AgeAtDeath optional The age at death (death events only, of course).
DateOfBirth optional The date of birth (death events only, replaces AgeAtDeath from 1969).
Mother optional Mother's (maiden) name (birth events only).
Spouse optional Spouse's surname (marriage events only).
FicheRange optional The range of surnames covered by the fiche this entry was on.
FicheNumber optional The number of the fiche this entry was on.

Adding The Info Header
The info header is added automatically to any worksheets exported as FreeBmd.
Use the .bmd file suffix for your FreeBMD files.

Editing The Info Header Details
To edit the info header, select ' **Tools' > 'View/Edit Info Headers' > 'FreeBMD Info** '.
If this menu option is unavailable, go to ' **Tools/Preferences** ' and under the ' **Export** ' tab select ' **Export FreeBMD** '.

The FreeBMD header Information
Email : This is the email address of the transcriber (optional).

Sequenced : is one of SEQUENCED, RANDOM or ONENAME.
SEQUENCED : Should be used to transcribe complete pages from the index.
RANDOM : Should be used to transcribe entries that are not related to the location in the index page.
ONENAME : Should be used to transcribe sections of pages that relate to a single name.
RecordType : BIRTHS, MARRIAGES, or DEATHS.
CharacterSet : The characters set you are using. (This is entered for you).

Each time a FreeBMD (.bmd) file is opened, the header information is retrieved (if it exists).
When a FreeBMD (.bmd) file is exported, the current header information is saved in the file.
It is quite acceptable to open a csv file, edit the FreeBMD header info, and then export as FreeBMD.

FreeREG Header Information

This is the "Header" information that will be included in the header of files exported as FreeREG. This is required by FreeREG to enable them to understand the data and who provided it.

Submitter Email e-mail address of the submitter.

Password Your password (it may be required in the future). At the present time simply use password for this field.

Record Type The type of register information being submitted. It can be BAPTISMS, MARRIAGES or BURIALS.

CharSet The character set to export as
Submitter Name The name of the person submitting the file.
Submitter County The group with which the submitter is working.
Transcribe Filename
The file name must consist of:
a three letter Chapman code for the county
a three letter placename code assigned by the person uploading the file
BA|BU|MA][part number, if any

Transcribe Date The date of transcription (dd-mon-yyyy)
Credit Name Enter the transcriber name. This is used to provide credit information for others involved in the transcription of the file.
Credit Email The transcriber e-mail address. This is used to provide credit information for others involved in the transcription of the file.
Modification Date The date this file was modified.
Comment1 It is suggested that this is used to detail the source of the data.
Comment2 Use this for any copyright restrictions, or any comments you wish to make about the file.

Add +LDS Select this if two extra fields for Film and Image number have been added to every data record.

The FreeREG header information will be included in the header of files exported as 'FreeREG' . The FreeREG file is saved with a '.frg' suffix.

You will need to change the file suffix to '.csv' before sending it to your coordinator.

Exporting to Xml (raw)

The Xml raw export creates a clean xml file that can be used by any application capable of reading xml. This includes your web browser.

How to export to xml
1. From the **'File** ' menu, choose ' **Export > Xml (raw)** '.
2. Choose the filename to use for the xml file.(default is the current filename + .xml)
3. Click the 'Open' button.

Gedcom Import

Importing a gedcom file.

1. Select ' **Open Document** ' from the ' **File Menu** ' or click on the open button in the worksheet toolbar.

2. From the ' **Open** ' dialog ' **Type** ' dropdown list, select ' **IMPORT Gedcom** '.

3. Select the gedcom file, and click ' **Open** '.

Columns not used by the imported gedcom file are hidden in the worksheet. These can be made visible via the ' **Column Control** ' if needed.

When embedded notes are encountered during import, they are saved to text files. A folder is created with the name of the gedcom file, appended with ' **_files** '. Links to these text files are placed in the notes column.

To view a notes file, right click on a cell in the ' **Notes** ' column and choose ' **View this link** '.

To edit a notes files, when in edit mode, right click and choose ' **Edit note** '.

Gedcom Export

Exporting a gedcom file.

```
1. Open the '**File**' menu and choose '**Export**' and then '**Gedcom**'.
2. Select the path and filename.
3. Click '**Save**'
```

Any "notes" files will be embedded into the gedcom file during export.

Not all worksheets can be saved as a gedcom file. The gedcom file will contain information about individuals, and the worksheet needs to have the correct information for it to be of any use as a gedcom.

The minimum column titles that genscriber will accept for a gedcom export, are **Person_Id** *, **Surname, Given_Name, Sex** .
The columns you want to export must have one of the titles that genscriber will recognise as a gedcom field.
The worksheet may contain other titles, but they will not be exported.

Note *: The Person_ID can be created automatically during export.

The table below lists column titles that are recognised as gedcom fields. The titles are not case sensitive and may contain spaces or underscores anywhere in the title. You can use any of these column titles in any order.
I have tried to add enough variation in the names to avoid making it to restrictive. If I have missed a commonly used name, please contact me.

GEDCOM_TAG	ACCEPTABLE COLUMNS TITLES
INDI	INDI, Individual, ID, Person, PersonId
SURN	SURN, Surname, Lastname
GIVN	GIVN, Given, GivenName, Givennames, FirstName, FirstNames, Forename, Forenames
SEX	SEX, Gender
BIRT (DATE)	BIRT, Birth, Born, DateBorn, BirthDate
BIRT (PLAC)	BirthPlace, PlaceBorn
BIRT SOUR)	BirthSource

BAPT (DATE)	BAPT, Baptism, BaptDate, BaptismDate
BAPT (PLAC)	BaptPlace, BaptismPlace
BAPT (SOUR)	BaptSource, BaptismSource
DEAT (DATE)	DEAT, Death, DeathDate
DEAT (PLAC)	DeathPlace
DEAT (SOUR)	DeathSource
BURI (DATE)	BURI, Burial, BurialDate
BURI (PLAC)	BurialPlace
BURI (SOUR)	BuriSource, BurialSource
OBJE (FILE)	Image, Photo, ImagePath, PhotoPath, MediaPath, ImageFilePath, PhotoFilePath, MediaFilePath, ImageLink, Media, MediaLink, PhotoLink
OCCU	OCCU , Occupation
SOUR	SOUR , Source
NOTE	NOTE, Notes
Father ID	Father , FatherID
Mother ID	Mother, MotherID
Spouse ID 1	Spouse , SpouseID
MARR (DATE)	Married , Marriage, MarriedDate , MarriageDate
MARR (PLAC)	MarriedPlace , MarriagePlace
MARR (SOUR)	MarriedSource , MarriageSource
MARR (STAT)	MarriedStatus , MarriageStatus, MarrStat, MarriageStat
Spouse ID 2	Spouse2 , SpouseID2
MARR (DATE)	Married2 , Marriage2, MarriedDate2 , MarriageDate2

MARR (PLAC)	MarriedPlace2 , MarriagePlace2
MARR (SOUR)	MarriedSource2 , MarriageSource2
MARR (STAT)	MarriedStatus2 , MarriageStatus2, MarrStat2, MarriageStat2

Spouse Id 3	Spouse3 , SpouseID3
MARR (DATE)	Married3 , Marriage3, MarriedDate3 , MarriageDate3
MARR (PLAC)	MarriedPlace3 , MarriagePlace3
MARR (SOUR)	MarriedSource3 , MarriageSource3
MARR (STAT)	MarriedStatus3 , MarriageStatus3, MarrStat3, MarriageStat3

Spouse Id 4	Spouse4 , SpouseID4
MARR (DATE)	Married4 , Marriage4, MarriedDate4 , MarriageDate4
MARR (PLAC)	MarriedPlace4 , MarriagePlace4
MARR (SOUR)	MarriedSource4 , MarriageSource4
MARR (STAT)	MarriedStatus4 , MarriageStatus4, MarrStat4, MarriageStat4

CENS (DATE)	CENS, Census, CensDate, CensusDate
CENS (PLAC)	CensPlace, CensusPlace
CENS (ADDR)	CensAddress, CensAddr, CensusAddr, CensusAddress
CENS (AGE)	CensAge, CensusAge
CENS (RELA)	CensRela, CensusRela, RelationshipToHead, Relationship, CensRelationship, CensusRelationship
CENS (NOTE)	CensNote

Pdf Export

Worksheet PDF Export.

Exporting large worksheets to pdf can be a problem. If the exported pdf pages are kept at A4 size, a large worksheet is either unreadable, or it has to be spread it over multiple pages.

I decided to make the pdf export resize the pages to fit the worksheet. This creates non-standard pages, but allows scrolling to view large worksheets. The only problem is if you want to print it, but I will leave that for you to ponder over.

Only visible fields are exported. Hiding fields is one way of reducing the width of the exported pages.

TextPad PDF Export

The textpad does not have any problems like the above.

Exporting Html

Html Export

This is one of my favourite features.The html export is very useful for providing someone with an easily readable version of your worksheet or even creating webpages.
The html files produced by genscriber are compliant with   W3C   standards and are suitable for use on the Internet.
I only cover the basics here. If you need to know more, please contact me and I will provide detailed instructions.

How to export to html
1. Set the html options in 'preferences'.

2. If there are columns you do not want to export, hide them.

3. If the 'Set fixed width' option is chosen, adjust the worksheet column widths to suit.

4. From the 'File' menu, choose 'Export > Html Files'.

5. Choose the basename to use for the html files.(default is the current filename)

6. Click the 'Open' button.

A folder with the name of your base filename choice will be created automatically, and the exported html files will be place in this folder.
If you chose to have column sorting, then a folder called 'tablesorter' will be created in the html export folder. This contains the jquery library and scripts.

Html Export options in Tools>Preferences

Html Rows Per Page
Each file will contain a maximum number of rows as per the setting of 'Html Rows Per Page'. The default is 200.
Each html file created will contain links to all other pages. The link will be the number of the html file (page).

Set fixed <column> widths
Sets fixed table and column widths when exporting html.
When this option is selected, the html column widths will be fixed according to the widths of the columns in the worksheet at the time of export.

If nowrap is also set, columns will expand to accomodate data.

Set nowrap in <td>
Stops text wrapping in cells when exporting html.
This is best set ON when 'Set fixed' above is OFF.

Use external css file
Instructs genscriber to create an external css file when exporting. Useful if you want to change the styles in one place.

Add table sort
Option to add scripts to sort the table data.

Use dd/mm/yy format for shortdate
The date format for shortDate is 'mm/dd/yy ' by default. To change to 'dd/mm/yy ' , select this option.

Add a webpage heading
This will add html above the exported tables. It will be added to every web page.
When you select this option, the 'Heading Html' below becomes active.

Heading Html
Add the html for your heading here, or if you prefer, edit the 'header.html' file directly.
The header.html file can be found in the '/include/web/' folder. I have added example html for navigation. Modify or delete this as you like.

Add a webpage footer
This will add html below the exported tables. It will be added to every web page.
When you select this option, the 'Footer Html' below becomes active.

Footer Html
Add the html for your footer here, or if you prefer, edit the 'footer.html' file directly. The footer.html file can be found in the '/include/web/' folder.

Location of html template files
The html template files can be found in the '   GenScriber/include/web/   ' folder.

header.html

This is the header. Modify this as you like. It is displayed in a div above the tables.

footer.html

This is the footer. Modify this as you like. It is displayed in a div below the tables.

tablestyle.css

This is the css file. Modify this to change the style of your web pages.

tablesorter

The folder contains the files that are included when you choose to have table sorting. The css file for tablesorter is located in this folder.

Special placeholders

There are a few placeholder that can be used in the header and footer templates.
{YEAR} This will be replace by the current year.
{FILETITLE} The name of the worksheet file.
{USERNAME} Your username used by the system.

I can add more of these later if they prove to be useful.
–

About the css files

The internal and external css are both derived from a file in the 'include/web' folder called 'tablestyle.css'. If you want to change the css, edit this file before you export.

You can of course change the css file in the exported html folder at any time, but keep in mind that if you export again to the same folder, the css file in the exported html folder will be overwritten by the one from the include folder.

There is also a css file in the 'tablesorter folder' (include/web/tablesorter/tablesorter.css). This file is specific to the tablesorter headers.

About the table sort files

The table sort files are stored in the 'tablesorter' folder, as follows…

asc.gif   image used in the thead for sorted ascending.
dec.gif   image used in the thead for sorted decending.
bg.gif   image used in the thead for unsorted.
jquery.js   jQuery JavaScript Library v1.4.2.
tablesorter.js   TableSorter 2.0 javascript.
tablesorter.css   TableSorter   CSS   .

If the 'table sort' option is selected, these files will be copied to the html export folder, and javascript code is written to the exported html files.

During export, genscriber will scan all cells of the first 1000 records to determine the type of data.
The sort type will be set for the content found. If mixed data is found, the sort type will be set by majority.

May 23 1890	
23 May 1890	isoDate
1890/5/23	
1890-5-23	

Monday, June 15, 2009 1:45 PM	usLongDate
May 23 1890 23	

05/23/1890 or 05-23-1890	
shortDate is mm/dd/yy by default.	shortDate
add {dateFormat: "dd/mm/yy"} to change to uk dates.	

123.123.123.123	ipAddress

07:25am   time
07:25pm

The date format for shortDate is ' **mm/dd/yy** ' by default.
To change to ' **dd/mm/yy** ' , select the option in 'Html Settings', or include the following code into the javascript of each html page.

```
$("#gstables").tablesorter({dateFormat: "dd/mm/yy"});
```