

Evaluating the Gramps Data Model with the Evidence and Conclusion Process

Tom Wetmore, November 17, 2010

This note evaluates the Gramps data model on how well users can use the model to support the Evidence and Conclusion model. This is part of an effort to design BetterGedcom, the specification of a format that can be used to record genealogical data in databases and external file archives, and enable complete sharing of data between genealogical programs through transport files.

The author of this report is not an experienced Gramps user. He downloaded the program, read its documentation, studied its data model, and then used the program to build databases that explored all aspects of Gramps needed to support the Evidence and Conclusion process. Key to this was regularly exporting Gramps databases to their XML archival files and studying the files in detail. This was the best way to come to a detailed understanding of the basic structure of Gramps records and their relationships.

The Evidence and Conclusion process (E&CP) has seven steps. Each one will be outlined below with an evaluation of how well the Gramps model supports it.

One. A researcher finds a source of information that contains evidence that mentions persons he may be interested in. He creates a record to document the source.

Gramps has both repository and source records that store information about sources and the repositories that hold them. Source records use a few fixed attributes such as title, author, abbreviation and publication information, and the user can add any number of other key/value attributes. Other records refer to source records with references can add more attributes defining such things as page numbers.

Conclusion: Gramps fully supports this step.

Two. The researcher creates records to document each item of evidence in the source that mentions the persons of interest.

Gramps does not have a record type specifically for evidence, but locations in the Gramps model can be found to store evidence. One reasonable location is as an attribute to the source reference between a Gramps event record and its source record. Since source references can have any number of user-defined attributes, users can easily devise conventions for recording evidence. Another reasonable location is as custom attributes directly in event records. A disadvantage of these two approaches is the redundancy that occurs if more than one event derives from the same evidence. A third location to store evidence would be in Gramps note records. This solves the redundancy issue, but note records are unstructured records that only contain a single string.

Conclusion: Gramps partially supports this step through user conventions.

Three. The researcher creates records to document each event mentioned in the evidence, where an event describes something that happened to one or more persons. Events occur at specific times and places, involve one or more persons as role players, and may serve to establish or change relationships between persons.

Four. The researcher creates records to document each person mentioned in each event, containing only information available from the evidence. The event and associated person records must be treated as a cohesive whole since the person records may hold information that is true only within the context of the event, e.g., the person's name or age or place of residence, at the time the event occurred.

Five. The researcher continues this process, completing steps 1 to 4, for a number of sources, until he has built up a number of groups of associated event and person records that contain all he has discovered about a set of persons.

Gramps has event records that hold the type, date, and place of events, and allows users to give them references to sources, notes, external files, URLs, and arbitrary attributes. However Gramps event records do not contain references to the person records that play roles in the event. Instead roles are defined in Gramps person records that have event references; event references refer to event records and have a role tag that indicates the role the person has in the event. However, users can create and remove event references at any time. In the E&CP it is important that the event and person records extracted from evidence be bound permanently together by their roles, since it is this combined set of these records that encodes each evidence-based event. This points out a problem Gramps has in supporting the E&CP – in Gramps events and persons are independent records that users choose to link together or not. In order to support the E&CP, users would have to follow strict conventions when creating events and persons from evidence, linking those records together through the persons' event references, that can never be removed or modified.

Conclusion: Gramps supports these steps but only through adherence to strict user conventions that have no support in the Gramps user interface.

Six. The researcher reasons about the available person records and sorts them into groups, where each group contains the person records that the researcher believes refer to a single real person. The researcher builds these groups based on experience and good practices, and records the justification for each grouping decision. Because later evidence may prove some groupings to be incorrect, the grouping operation cannot destroy or remove original event or person records. Groups need to be supported by the model.

Gramps has no group record and no user interface support for evaluating a set of person records for the purpose of partitioning them into groups that the user believes represent the same real persons. However, Gramps has a mechanism, called association, which allows person records to refer to other person records, which does provide a way to create person groups. Each association is a reference to another person record, and each association has a tag that specifies the association type. For example a Gramps user could use the tag *evidence* to associate a person

record to all the person records in one of the user's groups. However, the independence of event and person records pointed out above, and the fact that Gramps has no direct support for distinguishing evidence persons from conclusion persons, points out the fact that, even though Gramps supports the notion of evidence as well or better than most other programs, it is still fundamentally a system for expressing conclusions. Users of Gramps are not intended to keep multiple person records in their databases that represent the same real person. Person records are intended to refer to all of the events from all sources that the user believes involves to the same real person. As soon as two person records are discovered that the user believes are the same person he is encouraged to merge them together, creating one person from two that now holds the combined information and references from the original two. This is a nearly irreversible operation.

Conclusion: Through strict use of conventions involving the association mechanism, users can use Gramps to implement person records of both evidence and conclusion types. There is no internal support for this concept and no user interface that can aid the user in implementing the convention – Gramps is simply not intended to support evidence and conclusion persons. For example, Gramps has no knowledge that persons who are grouped together using an evidence association are thought to be the same person by the user. As far as reports or displays are concerned they are all separate persons.

Seven. The researcher reasons about the event records associated with the persons in each of the person groups in order to infer the relationships that existed between the real persons represented by the groups. The data model allows him to establish these relationships, possibly through new records that represent inferred genealogical events, or possibly by establishing relationship links between groups.

There is no need to go into detail here. In Gramps event records can come from evidence or they can come when users are entering information about persons they have already determined to exist. There is no mechanism in Gramps that allows users to group events together into higher level groups.

Overall conclusion. Gramps was not intended to support the Evidence and Conclusion process; there is nothing in the Gramps documentation that implies that the process was considered in the design of the Gramps model or user interface. However, Gramps is so flexible and has such a rich set of data types, that by strict adherence to conventions that are not explicitly supported by the Gramp user interface, users can implement most aspects of the process.