Snapshot of the page 11 Apr 2011 Created by Adrian

# Introduction

I have been concerned for some time that there is a gap in our analyses between the genealogical research process, the evidence & conclusion model and my aspirations for a "scientifically" robust process and write-up. This page is an attempt to fill in that gap and define more specifically what we need - in my *personal* opinion - to be documenting and therefore what I would like to see in my ideal genealogy research software. Ultimately though, the focus of the page is less on the process, and more on driving out what data items are involved in that process, so I can be more confident about what's needed in the BetterGEDCOM Data Model.

As with all my process descriptions, I usually don't mention whether or not software will be used at any point, because a process is about "what" is done, not "how". Where I do mention software, it's because avoiding doing so makes the words harder to read.

It may very well be that this page does no more than document what you, the reader, felt was obvious. If so fine, but I have satisfied myself that I have filled in the gap.

This process researches a specific event, attribute or relationship concerning something or someone. It does not attempt to set an overall strategic direction.
(*Note - I have not yet tested this process in my mind against large-scale family reconstruction, but suspect the same steps appear - just more often.*)

1. **Set and record a focussed goal** (e.g. "Who were the parents of X?") (*c.f. Tom Jones, "Inferential Genealogy" course handout, 2010, Family Search*)
*Input*

- Current conclusions about people, things, etc

*Output*

- Focussed goal – what things or people are to be investigated? What relationships, events or attributes do we want to know about them?

2. **Create a "reasonably exhaustive" initial search plan with specific objective(s)** about what to look for in order to reach the goal. This plan starts as an initial plan - it is highly likely that it will be necessary to loop back here and create a revised plan later on in the light of information found - "No plan survives first contact" (Field-Marshal Helmuth von Moltke the Elder). Indeed, the initial plan may be only half a plan, with the rest being defined only in the light of the first set of discoveries. The search plan needs to

be broad enough in time, space and people to trap potentially useful information (c.f. Tom Jones, also the Genealogical Proof Standard.) Do not be afraid to include speculative items, e.g. "Anything in Chester Quarter Sessions records for the 1820s?".
*Input*

- Focussed goal (from previous step).
- Current conclusions about people, things, etc
- Useful data about what sources to look for when looking for evidence relating to specific attribute / event values (e.g. "parents' names can be found on Scots death certificates") (probably this data is held elsewhere)
- Useful data about who might have created those sources (e.g. "post-1837 marriage certificates in England were created by register offices and / or churches") (probably this data is held elsewhere)
- Useful data about those sources e.g. "pre-1837 marriages in England were always in Church of England parish churches" (probably this data is held elsewhere)
- Where to look for those potential sources (e.g. list of where Cheshire parish registers can be found) (probably this data is held elsewhere)

*Output*

- List of specific objectives ("what"), each with a statement of "how" I intend to fulfil each objective – search criteria *and* intended logic (e.g. "Look for all marriages in Nantwich, Acton and Wistaston between 1820 and 1840 with groom's name = X *and* check matches to original couple")
- Location of potentially *relevant* sources ("where") needed to fulfil the objectives above, forming an initial search log (e.g. "Microfilms of parish registers (PRs) for Nantwich, Acton and Wistaston, at Chester Archives". This log will then be completed in the next steps to summarise what has been found.) (*The list of objectives and the initial log, taken together, make up a search plan*)
- Documented assumptions – "As both X and Y live close to where they were born, according to the censuses, it is assumed that they were married reasonably close to there." Also "It is assumed that neither of N and W were Quaker or Jewish" – these are the two exceptions to the rule about marriage in CofE church.

3. **Carry out research** according to the current search plan. (*When searching for paper-based records, this step takes place in Archives, Record Offices, etc. When searching internet based records, this takes place at a computer terminal and the division between this and subsequent steps might tend to disappear.*). (There are personal decisions to be made about how much to record for sources that are close to the search criteria but do not match - if it's been a long journey and you won't be back for a while, it might be tempting to record "close" sources in case they turn out to be for a relative.)
*Input*

- List of specific objectives, each with a statement of "how" I intend to fulfil each objective (from previous step)
- Location of potentially *relevant* sources needed to fulfil the objectives above, forming an initial

search log (from previous step)
- Documented assumptions (from previous step)

*Output*

- Contents of the researched sources, documented in such a way as to be understandable (e.g. a series of marriage transcripts for people satisfying the search criteria) and with enough data to enable accurate citations, record provenance, etc.
- Updated search log saying for each source, what's been searched, what was missing, etc.

3.5 **Check your understanding** of the records that have been judged to have useful information. Understanding why a record was created will help interpret the information in it (*For instance, does the grant of probate say 'Personal Effects' and / or 'Real Estate' - do you understand the difference?*). (*c.f. Tom Jones, "Inferential Genealogy"*)

4. Assemble the research. Assess the quality of the source material. Look for any patterns, matches, differences, etc. that might be meaningful. (*These are not meant to be sequential steps but can take place in parallel*)
Extract the evidence - i.e. the information that is relevant to the specific objective(s).
Can you demonstrate ("prove") that the person referred to in the evidence is the one that the objective(s) needs?
**Analyse the evidence** to see if each objective has been met. Document that analysis.
Are there any interim conclusions? Is there any conflicting evidence? (*e.g. this looks like him but it's the wrong father*) Any partial progress? (*e.g. These are the marriages matching our couple but there is more than 1 match, so we cannot yet tell which is their marriage*). Any evidence that contradicts any hypothesis? (*e.g. Implication of complete search is that they were not married after all*)
*Input*

- List of specific objectives, each with a statement of "how" I intend to fulfil each objective (from previous steps)
- Contents of the researched sources, documented in such a way as to be understandable and with enough data to enable accurate citations, record provenance, etc. (from previous step).
- Updated search log saying for each source, what's been searched, what was missing, etc (from previous step)
- An understanding of the nature of the researched sources (probably held elsewhere)

*Output*

- Evidence relevant to the objective(s), including evidence of identity
- Results of analysis
- Interim conclusions (including identity) – if any
- Conflicting evidence – if any

5. If the specific objectives have not been met, **or** if there is conflicting evidence that cannot be resolved, **return to create new search plan** with revised specific objectives (*Some conflicts can be accepted if they can be resolved, i.e. explained away in a plausible manner - e.g. it was 50 years after the marriage and the son giving the information was born long after that marriage*.)

6. If the specific objectives *have* been met with **no** unresolved conflicting evidence, analyse results to see if focussed goal has been met. Is there any remaining unresolved conflicting evidence at this higher level when reviewing at this higher level? Document the analysis. What are the conclusions? **Record the analysis and conclusions in some form of proof summary or proof argument**.
*Input*

- As outputs from above

*Output*

- Proof (summary or argument) giving results of analysis and final conclusions – if any (See BCG for samples).
- Conflicting evidence – if any

7. If there is any unresolved conflicting evidence, **return to create new search plan** with revised specific objectives.

8. **Enter the final conclusions into the genealogy application**, either

- merging these conclusions into existing people, events, objects, etc, deleting or replacing conclusions that are no longer accepted (i.e. use the "conclusion-only" model) or
- creating new people, events, objects containing the new conclusions plus old conclusions that are still accepted (leaving the old detail there but marked up as "superseded") (i.e. use the "evidence conclusion model")
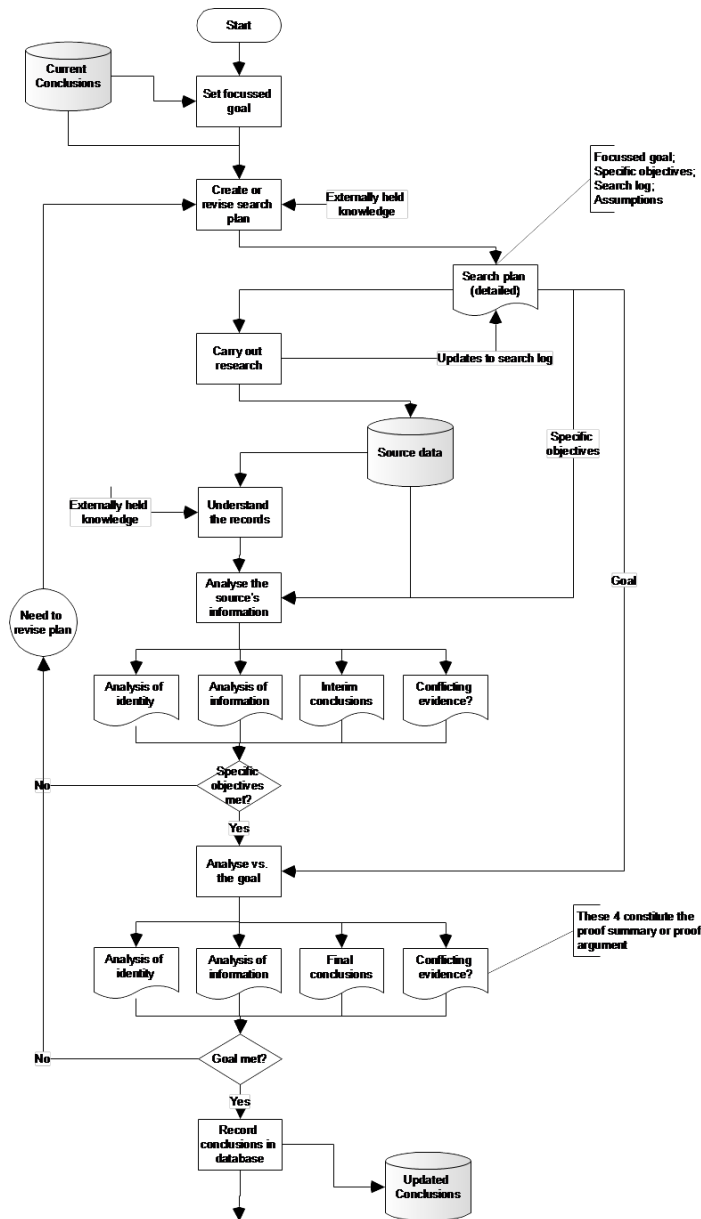
## Caveats

I talk of "proof" and "final conclusions". It is unlikely that proof in a complex genealogical study can reach the standard of proof required in a criminal case ("beyond a reasonable doubt") - the Genealogical Proof Standard exists to provide criteria to judge the standard of proof obtained. Nor is any conclusion really final as the appearance of previously unsuspected information may throw everything into suspicion.

## Application Software

The crux of the matter is this - what do I want to see in an application? And therefore in the BetterGEDCOM Data Model? The answer is - *everything* that's recorded as an input or an output above.

Actually, there are exceptions denoted above by the phrase "probably this data is held elsewhere" since otherwise we'd end up dumping all the text books into our applications. But after these first steps, I'd really like all that lot to go into my application.

# Diagrammatically



# Data Analysis

First cut list of entities - excluding those already clearly covered by GEDCOM - i.e. persons, families, etc. This list is somewhat descriptive, rather than specific.

- Focussed goal – *what things or people are to be investigated? What relationships, events or*

*attributes do we want to know about them*?
- Search Plan -
    - List of specific objectives ("*what*"), each with a statement of "*how*" I intend to fulfil each objective – search criteria and intended logic
    - Initial search log containing location(s) of potentially relevant sources ("*where*") needed to fulfil the objectives above. *This log is later updated completed to summarise what has been or cannot be found where.*
    - Documented assumptions
- Contents of the researched sources, documented in such a way as to be understandable, and with enough data to enable accurate citations, record provenance, etc (*as per GEDCOM Source-entity but with extra attributes compared to current GEDCOM*?)
- Progress against specific objective
    - Evidence, including evidence of identity
    - Results of analysis
    - Interim conclusions (including conclusions re identity) if any
    - Conflicting evidence if any
- Proof argument or summary made up of
    - Evidence, including evidence of identity
    - Results of analysis against goal
    - Interim conclusions (including conclusions re identity) against goal if any
    - Conflicting evidence against goal if any

The indented bullets are intended to imply a probable relationship - e.g. the higher level bullet consists of the lower level ones.

Comments?