USING A (LINEAGE LINKED) FAMILY PERSPECTIVE OVER HOUSEHOLD TO IMPROVE RECORD LINKAGE SUCCESS WITH CENSUS (AND OTHER) DATA COLLECTIONS

Prepared by David S. Barss, AG®
Historical Family Reconstitution
FamilySearch
12 Apr 2010

Creating lineage linked families, and in some cases pedigrees, from census data provides a broader foot print than can be obtained by using the household perspective and increases the opportunities for record linkage success when matching and merging the census data with other records.

Traditional data gathering and record linkage activity using census data has focused on the household as the main unit of identification. By using lineage linked families as the main unit of identification more data can be accumulated from the census which will improve opportunities for linking records across the census years or with other data sets.

There are often data elements within the census, that when used with stated relationships, will help to identify and capture broader family connections, especially if the data is preserved in a lineage linked format. If the census includes a marriage date it expands these opportunities even further. Interpreting the census data can:

- expand the breadth of the records used in matching
- preserve linkage that is given in the census without having to recreate it
- capture sibling families where no parent is present
- capture multi-generational families found within the household
- capture “hidden families” that are not directly identified by stated relationships
- provide a more accurate representation of some families
- eliminate the need to deal with changing relationship roles of an individual as they are tracked across census years
- provide, based on local custom, a father’s name that is not stated in the census
- capture family data for families not related to the head of the household like servants, boarders, or laborers

These things are accomplished by the use of relationship pointers, and in some cases, the creation of a “derived record” based on census data used to link the family together.

Converting Census data to Lineage Linked Family Data

We are using a data conversion tool called “CensusToGed” that was created for us by Pleiades Software Development which we continue to revise and improve as we use it. This data converter takes the census data from a delimited text file to a Gedcom file.

We have adopted the data coding, and code values, that are used by the North Atlantic Population Project at the Minnesota Population Center at the University of Minnesota in

Minneapolis. These codes and their associated values can be seen at the website for the North Atlantic Population Project (www.nappdata.org). This coding process identifies:

- Each household
- The head of the household
- The relationship of each person within the household to the head
- The position of each person within the household
- The position of each person's spouse within the household
- The position of each person's father within the household
- The position of each person's mother within the household

We have added to their coding system a few values that allow us to better handle Sibling-in-law relationships. For Norway data we have also added to the census data a column that identifies what the individual's father's given name would be based on the presence of a patronymic surname.

For the Norway data using the relationship codes, **without** the position pointers, the CensusToGed data converter can capture the following simple relationships:

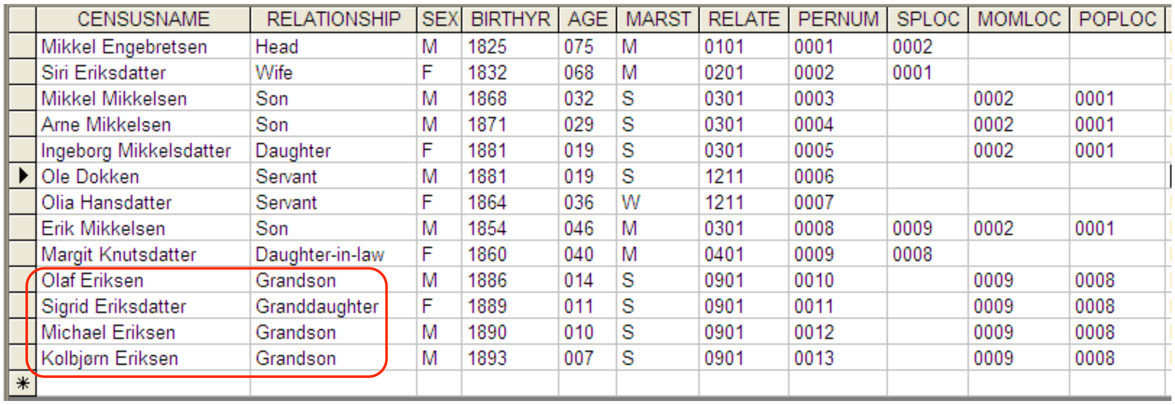
- Head
- Spouse
- Child
- Sibling (including sibling families with a derived father)
- Parent
- Parent-in-law
- Sibling-in-law
- Derived father's given name based on patronymic name patterns

Using the relationship codes **and** the position pointers we can capture all of the above relationships plus others that are more complicated:

- Child-in-law
- Grandchild (if a parent is within the household)
- Sibling's spouse (and their families if present)
- Sibling-in-law's spouse (and their families if present)
- Non-relative Spouses (and their families if present) – these are servants etc.

An example family from the 1900 Norway Census of Sør-Aurdal, Oppland, Norway will illustrate what the CensusToGed converter can do in both cases. The census extract with the coding (RELATE, PERNUM, SPLOC, MOMLOC, and POPLOC) is listed in Figure 1. The family as captured using the relationship codes and the position pointers is seen in Figure 2, and the same family as captured using only the relationship codes is seen in Figure 3. The significant difference between these two figures is the connection to the grand children who are marked with a red outline.

Figure 1 - 1900 Norway Census for Sør-Aurdal, Oppland, Norway – sample study household



	CENSUSNAME	RELATIONSHIP	SEX	BIRTHYR	AGE	MARST	RELATE	PERNUM	SPLOC	MOMLOC	POPLOC
	Mikkel Engebretsen	Head	M	1825	075	M	0101	0001	0002		
	Siri Eriksdatter	Wife	F	1832	068	M	0201	0002	0001		
	Mikkel Mikkelsen	Son	M	1868	032	S	0301	0003		0002	0001
	Arne Mikkelsen	Son	M	1871	029	S	0301	0004		0002	0001
	Ingeborg Mikkelsdatter	Daughter	F	1881	019	S	0301	0005		0002	0001
▶	Ole Dokken	Servant	M	1881	019	S	1211	0006			
	Olia Hansdatter	Servant	F	1864	036	W	1211	0007			
	Erik Mikkelsen	Son	M	1854	046	M	0301	0008	0009	0002	0001
	Margit Knutsdatter	Daughter-in-law	F	1860	040	M	0401	0009	0008		
	Olaf Eriksen	Grandson	M	1886	014	S	0901	0010		0009	0008
	Sigrid Eriksdatter	Granddaughter	F	1889	011	S	0901	0011		0009	0008
	Michael Eriksen	Grandson	M	1890	010	S	0901	0012		0009	0008
	Kolbjørn Eriksen	Grandson	M	1893	007	S	0901	0013		0009	0008
*											

Figure 2 - Head of household and descendants – with the grandchildren linked:

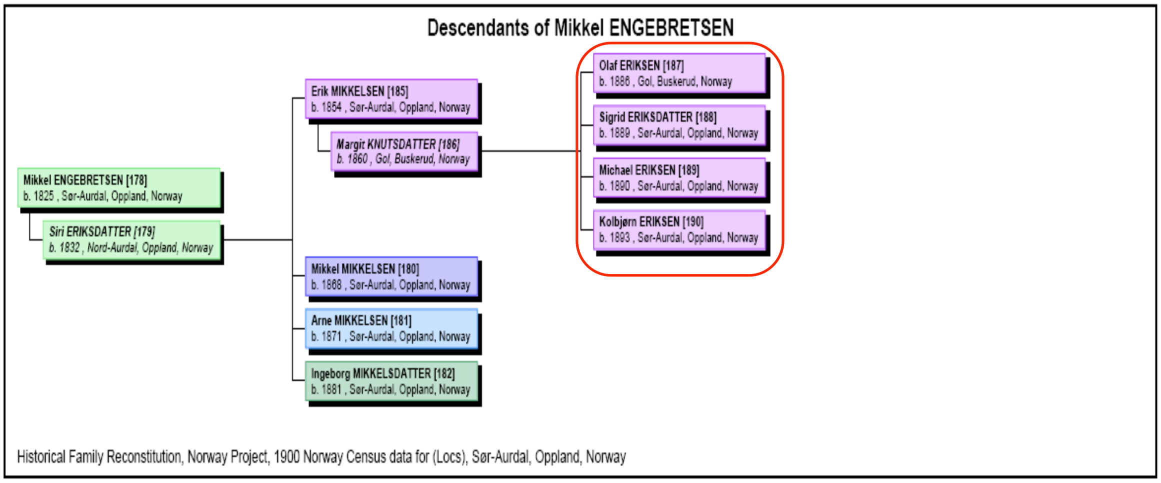
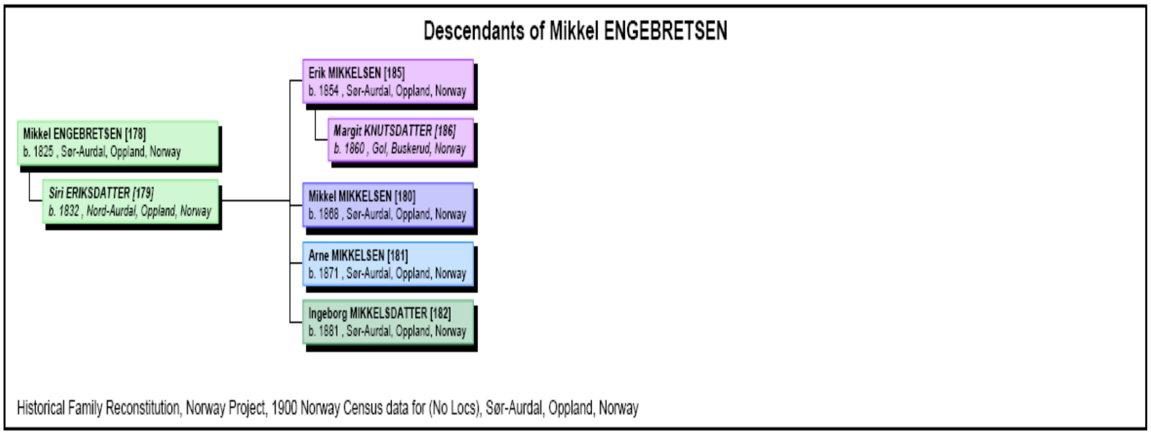
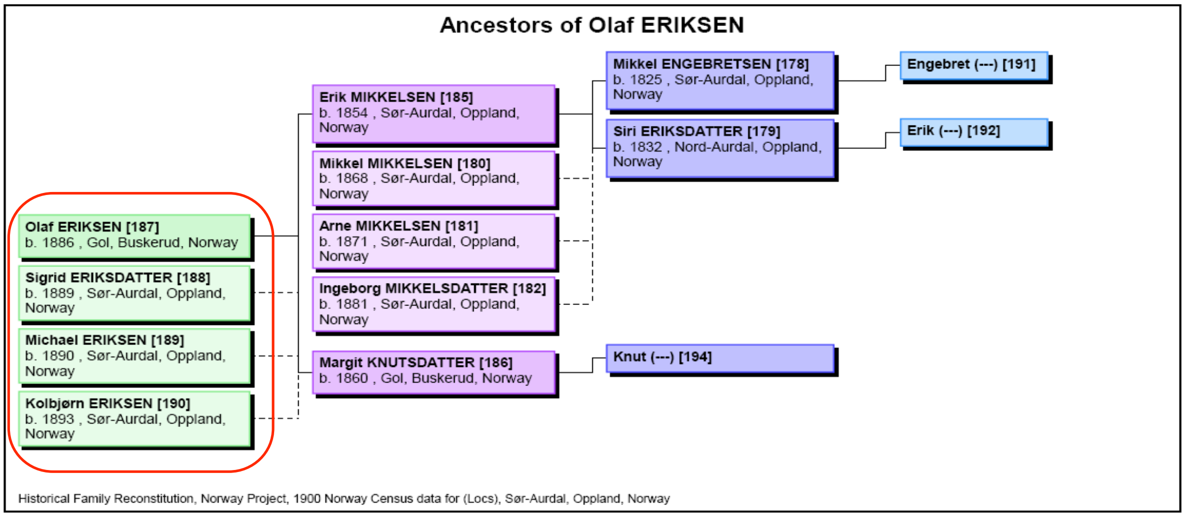

Descendants of Mikkel ENGEBRETSEN
Mikkel ENGEBRETSEN [178]
b. 1825 , Sør-Aurdal, Oppland, Norway
Siri ERIKSDATTER [179]
b. 1832 , Nord-Aurdal, Oppland, Norway
Erik MIKKELSEN [185]
b. 1854 , Sør-Aurdal, Oppland, Norway
Margit KNUTSDATTER [186]
b. 1860 , Gol, Buskerud, Norway
Olaf ERIKSEN [187]
b. 1886 , Gol, Buskerud, Norway
Sigrid ERIKSDATTER [188]
b. 1889 , Sør-Aurdal, Oppland, Norway
Michael ERIKSEN [189]
b. 1890 , Sør-Aurdal, Oppland, Norway
Kolbjørn ERIKSEN [190]
b. 1893 , Sør-Aurdal, Oppland, Norway
Mikkel MIKKELSEN [180]
b. 1868 , Sør-Aurdal, Oppland, Norway
Arne MIKKELSEN [181]
b. 1871 , Sør-Aurdal, Oppland, Norway
Ingeborg MIKKELSDATTER [182]
b. 1881 , Sør-Aurdal, Oppland, Norway
Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (Locs), Sør-Aurdal, Oppland, Norway

Figure 3 - Head of household and descendants – without the grandchildren linked:


Descendants of Mikkel ENGEBRETSEN
Mikkel ENGEBRETSEN [178]
b. 1825 , Sør-Aurdal, Oppland, Norway
Siri ERIKSDATTER [179]
b. 1832 , Nord-Aurdal, Oppland, Norway
Erik MIKKELSEN [185]
b. 1854 , Sør-Aurdal, Oppland, Norway
Margit KNUTSDATTER [186]
b. 1860 , Gol, Buskerud, Norway
Mikkel MIKKELSEN [180]
b. 1868 , Sør-Aurdal, Oppland, Norway
Arne MIKKELSEN [181]
b. 1871 , Sør-Aurdal, Oppland, Norway
Ingeborg MIKKELSDATTER [182]
b. 1881 , Sør-Aurdal, Oppland, Norway
Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (No Locs), Sør-Aurdal, Oppland, Norway

The significant of this lineage linked representation of the family and the foot print created by data is better observed from the grandchild's point of view. Using the first of the four grandchildren as our example, Figure 4 shows his pedigree with siblings where the position pointers were used, and Figure 5 shows what happens using the relationship codes alone which can not link the grandchildren to the family. You will also see in both of these examples, the addition of a father's given name as derived from the individual's patronymic surname. The grandchildren in these examples are again outlined in red.

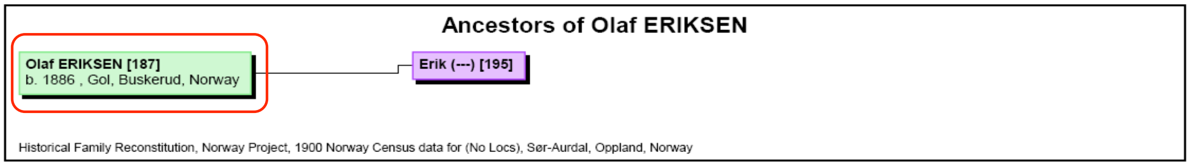
Figure 4 – Grandchild pedigree with siblings – where linked to the family


Ancestors of Olaf ERIKSEN

Olaf ERIKSEN [187]
b. 1886 , Gol, Buskerud, Norway
Sigrid ERIKSDATTER [188]
b. 1889 , Sør-Aurdal, Oppland, Norway
Michael ERIKSEN [189]
b. 1890 , Sør-Aurdal, Oppland, Norway
Kolbjørn ERIKSEN [190]
b. 1893 , Sør-Aurdal, Oppland, Norway
Erik MIKKELSEN [185]
b. 1854 , Sør-Aurdal, Oppland, Norway
Mikkel MIKKELSEN [180]
b. 1868 , Sør-Aurdal, Oppland, Norway
Arne MIKKELSEN [181]
b. 1871 , Sør-Aurdal, Oppland, Norway
Ingeborg MIKKELSDATTER [182]
b. 1881 , Sør-Aurdal, Oppland, Norway
Margit KNUTSDATTER [186]
b. 1860 , Gol, Buskerud, Norway
Mikkel ENGEBRETSEN [178]
b. 1825 , Sør-Aurdal, Oppland, Norway
Siri ERIKSDATTER [179]
b. 1832 , Nord-Aurdal, Oppland, Norway
Knut (---) [194]
Engebret (---) [191]
Erik (---) [192]

Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (Locs), Sør-Aurdal, Oppland, Norway

Figure 5 – Grandchild pedigree with siblings – where not linked to the family


Ancestors of Olaf ERIKSEN

Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (No Locs), Sør-Aurdal, Oppland, Norway

It is obvious from these two examples that the difference in the foot print created when using the position pointers is worth the effort needed to complete the additional coding. We will give other examples of what can be done to capture family data using the family as the main unit of identification in the census (over the household) in an appendix at the end of this paper.

Testing record linkage success

The 1900 Norway Census data for the Sør-Aurdal Clerical District in Oppland County, Norway was used to demonstrate the increased record linkage success that is achieved when using lineage linked family data over household or unlinked data. The 1900 Census data will be merged with the Bygdebok data for the same place.

The Bygdebok is a local genealogy/history that contains extended family lineages. The Bygdebok data for Sør-Aurdal has been extracted into a lineage linked database that was used as the second data component for this test.

We focused our first test on grandchildren in the census. They were a small group so we could examine all cases, and using the CensusToGed converter we could easily produce a linked and unlinked set of data for the 1900 Census.

There were 38 grandchildren listed in the 1900 Census for Sør-Aurdal. Of that group 6 had no parents in the household so they would remain as unlinked individuals when we converted the census data to family data. These 6 were not used as part of our test group. There were also 8 grandchildren that had parents in the census data, but they were not found in the Bygdebok data, which meant we would not find matches for them, so they were not used as part of our test group either. This left 24 grandchildren, found in 10 different households, as our test group. Each of these entries had a parent present in the census which created a three generation presence in the household.

Results of the merging for grandchildren example

Bygdebok and 1900 Census **with** linked grandchildren (Data set 1)

- 18 of 24 grandchild matches were found (75% matching success)

Bygdebok and 1900 Census **without** linked grandchildren (Data set 2)

- 0 of 24 grandchild matches were found (0% matching success)

The merging success of the linked grandchildren was lower than we had hoped for at 75%, but it was much better than the 0% obtained by the unlinked sample. Given the foot print examples above these results were not unexpected.

Results of merging the full 1900 Census data set with the Bygdebok data

Our two 1900 Census data sets used above were both lineage linked for most relationships outside of the grandchildren, so we could not use them as linked and unlinked data samples to test matching results across the whole census. To compensate for that we, took our full linkage data set (Data set 1) and removed all of the family pointers from the Gedcom file. The result was a list of all of the individuals in the 1900 Census data set without any family linkage. We used this as our unlinked data set.

Bygdebok to 1900 Census full data sets

- Data set 1 (with linkage) found 3414 people in 1659 clusters
- Data set 1 (with **no** linkage) found 98 people in 49 clusters
- Only 3% of the matches that were found **with** the lineage linked data were found in the data set **with no** linkage
- Even though the bygdebok data is fairly complete, without the corresponding family linkage in the census data, the majority of the matches could not be found.

Conclusion:

These test results show that for the grandchildren in the 10 household study group, using the lineage linked family data was critical for matching success. The merging software found 75% of the grandchildren that had links to their parents, siblings, and grandparents, where **none** of them were found without the family linkage. The test results across the full 1900 Census data set had only slightly higher success. Only 3% of the matches found in the lineage linked sample were found in the unlinked data set.

The conclusion from these tests seem clear to us. There is a lot of record linking advantage to be gained by converting the source data (in this case census records) to lineage linked family data, before merging it with other data collections for the same locality.

Appendix – Further examples of family relationships that can be captured using a lineage linked family perspective with census and other data sets.

This appendix will discuss and give further examples of how the census data was converted from households to lineage linked families. It will also illustrate some of the additional data that can be pulled from the census when it is converted to a lineage linked family view using both information that is in the census and what is implied by that census data.

In converting the extracted census data to a lineage linked database additional data fields were added to the original extraction. This included an identifier for each household (SERIAL), a code for the individual's relationship to the head of the household (RELATE), the position of the person within that household (PERNUM), and pointers that use those positions to identify who the individual's spouse (SPLOC), father (POPLOC), and mother (MOMLOC) are within that household. The relationship data within the census and these pointers allow the creation of the lineage linked families. These variables, and code values, have been adopted from the North Atlantic Population Project, at the Minnesota Population Center, University of Minnesota at Minneapolis (www.nappdata.org). We have added a few codes to better handle Sibling-in-law relationships, and data types other than census with the same conversion tool.

In the case of countries where the patronymic naming pattern is used, the father's first name can be derive from the child's patronymic surname, a data field was added to the census extract to capture these names.

In countries where the census recorded women with their married name, if one of the wife's parents, male siblings, or unmarried female siblings were in the household, her maiden name can be determined. A data field was also added to these census extracts to allow the capture of both the women's maiden and married names.

In the case where there is a marriage date (or year) present in the census data, it gives even more abilities to pull additional data from the census, such as children from the father's and/or mother's previous marriage (hidden families) and so on.

In some cases family relationships that are clearly stated, or even implied, could easily be preserved by creating a "derived record". These derived records are clearly identified in the final data set with an indication that they have been derived from relationship data in the census. These derived records serve a vital role in helping to bind these families together.

In all cases, the census extract for the household is captured as part of the source data for the head of the household, and all other family members point to their record for the full census extract. For servants, laborers, and others that are not related to the head of the household, the full census extract is also attached to their record. A strategic choice was made regarding households with more than 10 non-relatives in them. The records of the non-relatives point to the head of the household's record for the full census extract. This would include households that are schools, hotels, boarding houses, prisons, and so on. This was simply a space saving measure.

We will be using census extracts from the 1900 Norway census and the 1900 US Census to demonstrate what data can be captured from the census using the family perspective. We will try to present the view of the family that will show the most linkage, so not all views will be the same. In some cases we will use pedigrees, in others a descendency will be the best view, etc.

The 1900 Norway Census (Sør-Aurdal, Oppland, Norway)

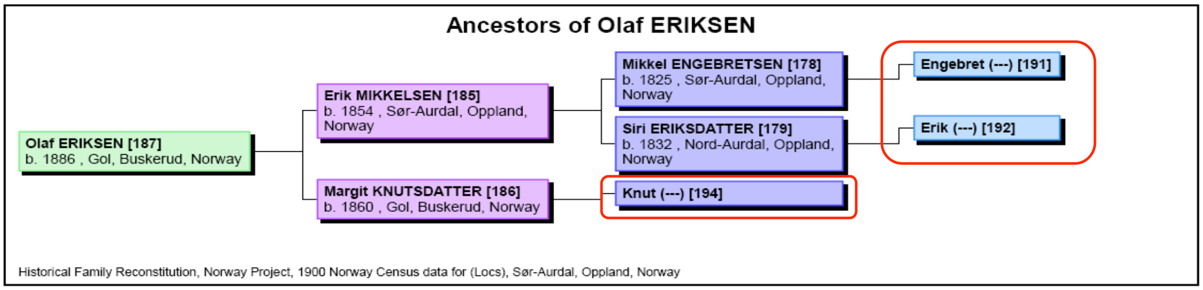
- Uses patronymic name patterns
- Records married women with their maiden names

The 1900 US Census (Lewis County, Washington, United States)

- Includes birth place of the individual and their parents
- Includes marriage year
- Records married women with their married names

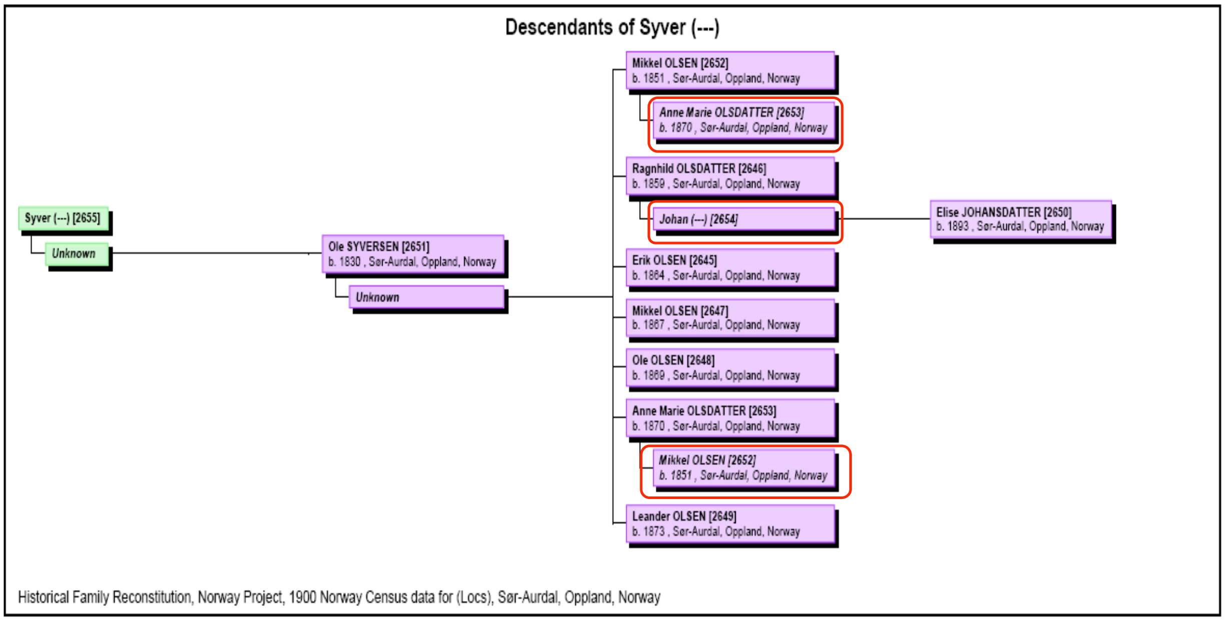
Examples of census households converted to lineage linked family data

Figure 6 - Norway Census – multiple generations, including the father's given name which has been derived from the child's patronymic surname


Ancestors of Olaf ERIKSEN
Olaf ERIKSEN [187]
b. 1886 , Gol, Buskerud, Norway
Erik MIKKELSEN [185]
b. 1854 , Sør-Aurdal, Oppland, Norway
Margit KNUTSDATTER [186]
b. 1860 , Gol, Buskerud, Norway
Mikkel ENGEBRETSEN [178]
b. 1825 , Sør-Aurdal, Oppland, Norway
Siri ERIKSDATTER [179]
b. 1832 , Nord-Aurdal, Oppland, Norway
Knut (---) [194]
Engebret (---) [191]
Erik (---) [192]

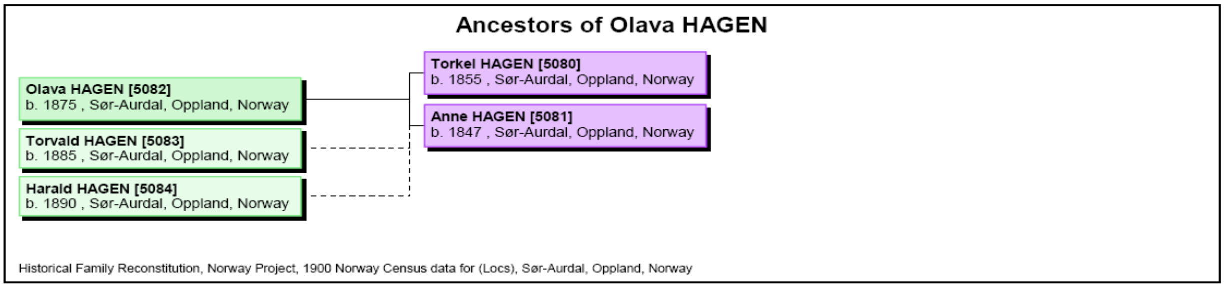
Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (Locs), Sør-Aurdal, Oppland, Norway

Figure 7 - Norway Census – children-in-law are correctly attached to the family, including one whose name was derived from his child's patronymic surname


Descendants of Syver (---)
Syver (---) [2655]
Unknown
Ole SYVERSEN [2651]
b. 1830 , Sør-Aurdal, Oppland, Norway
Unknown
Mikkel OLSEN [2652]
b. 1851 , Sør-Aurdal, Oppland, Norway
Anne Marie OLSDATTER [2653]
b. 1870 , Sør-Aurdal, Oppland, Norway
Ragnhild OLSDATTER [2646]
b. 1859 , Sør-Aurdal, Oppland, Norway
Johan (---) [2654]
Elise JOHANSDATTER [2650]
b. 1893 , Sør-Aurdal, Oppland, Norway
Erik OLSEN [2645]
b. 1864 , Sør-Aurdal, Oppland, Norway
Mikkel OLSEN [2647]
b. 1867 , Sør-Aurdal, Oppland, Norway
Ole OLSEN [2648]
b. 1869 , Sør-Aurdal, Oppland, Norway
Anne Marie OLSDATTER [2653]
b. 1870 , Sør-Aurdal, Oppland, Norway
Mikkel OLSEN [2652]
b. 1851 , Sør-Aurdal, Oppland, Norway
Leander OLSEN [2649]
b. 1873 , Sør-Aurdal, Oppland, Norway

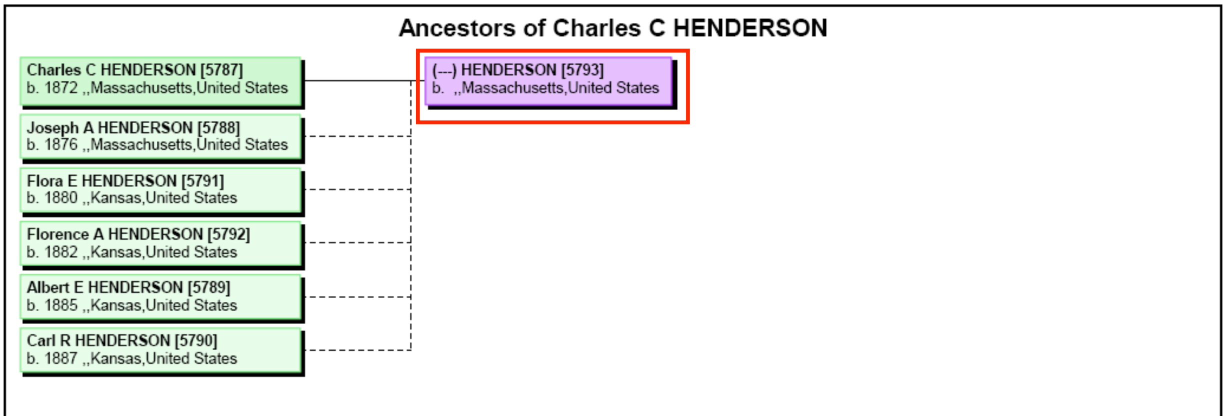
Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (Locs), Sør-Aurdal, Oppland, Norway

Figure 8 - Norway Census – family using non-patronymic surname, we did not create a derived father from the surnames


Ancestors of Olava HAGEN
Olava HAGEN [5082]
b. 1875 , Sør-Aurdal, Oppland, Norway
Torvald HAGEN [5083]
b. 1885 , Sør-Aurdal, Oppland, Norway
Harald HAGEN [5084]
b. 1890 , Sør-Aurdal, Oppland, Norway
Torkel HAGEN [5080]
b. 1855 , Sør-Aurdal, Oppland, Norway
Anne HAGEN [5081]
b. 1847 , Sør-Aurdal, Oppland, Norway

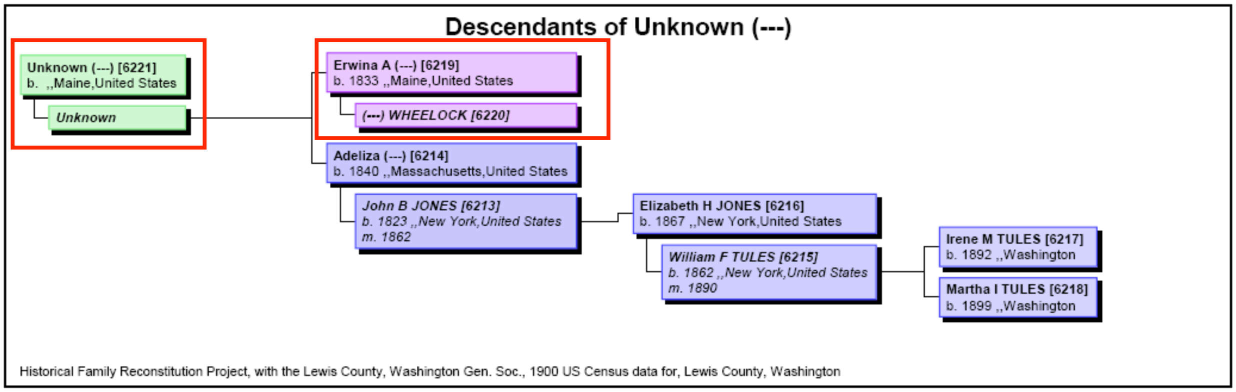
Historical Family Reconstitution, Norway Project, 1900 Norway Census data for (Locs), Sør-Aurdal, Oppland, Norway

Figure 9 - US Census – sibling family, where no parent was in the household, was preserved as a family by creating a derived father record. Father's birth place came from his child's record.


Ancestors of Charles C HENDERSON
Charles C HENDERSON [5787]
b. 1872 ,,Massachusetts,United States
Joseph A HENDERSON [5788]
b. 1876 ,,Massachusetts,United States
Flora E HENDERSON [5791]
b. 1880 ,,Kansas,United States
Florence A HENDERSON [5792]
b. 1882 ,,Kansas,United States
Albert E HENDERSON [5789]
b. 1885 ,,Kansas,United States
Carl R HENDERSON [5790]
b. 1887 ,,Kansas,United States
(---) HENDERSON [5793]
b. ,,Massachusetts,United States

Historical Family Reconstitution Project, with the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 10 - US Census – sibling-in-law (1), wife's sister is in the household, family preserved by creating a derived record with no name for their father. John B. JONES is the head of this household.



Historical Family Reconstitution Project, with the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 11 - US Census – sibling-in-law (2), siblings in the household are linked through a derived record for their father, the brother's wife is correctly identified and linked to her husband. Different codes are used for the two types of siblings-in-law. Gustaf SALTZER is the head of the household.

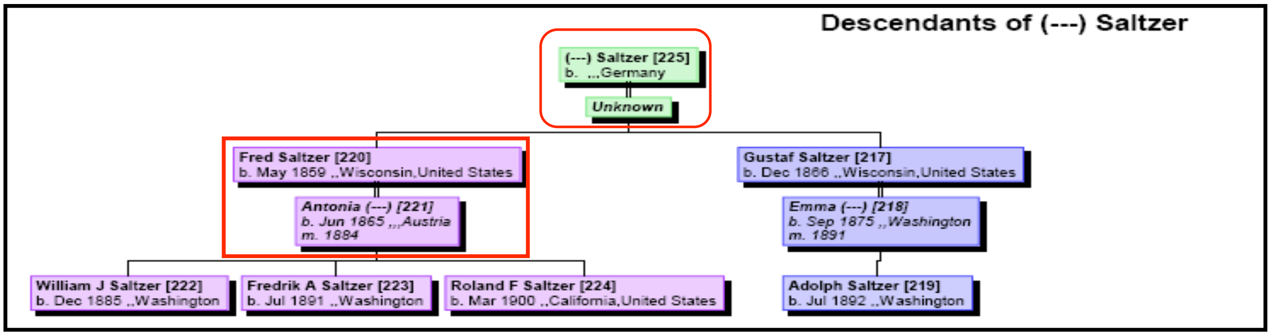
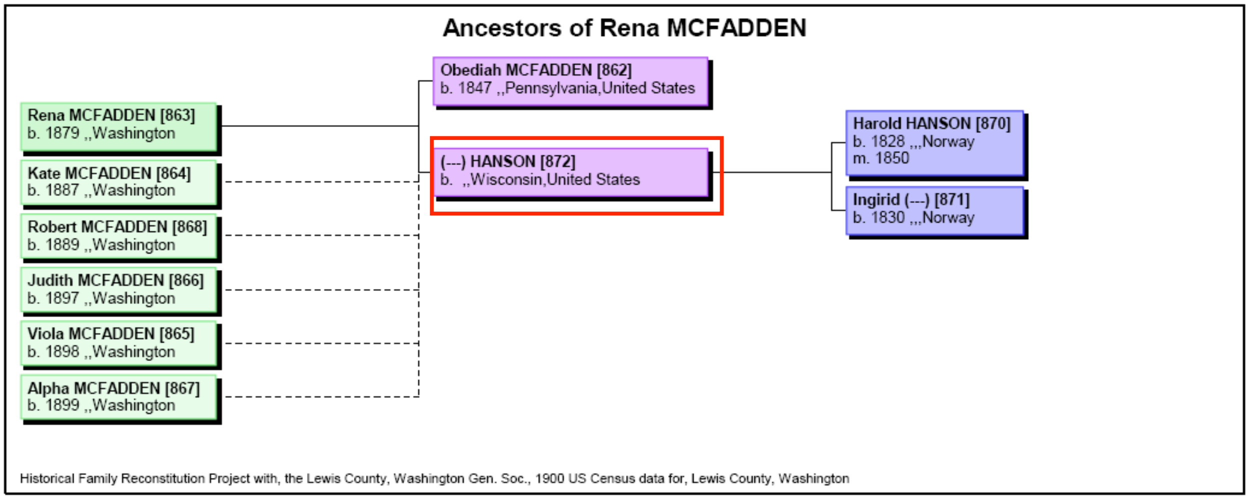
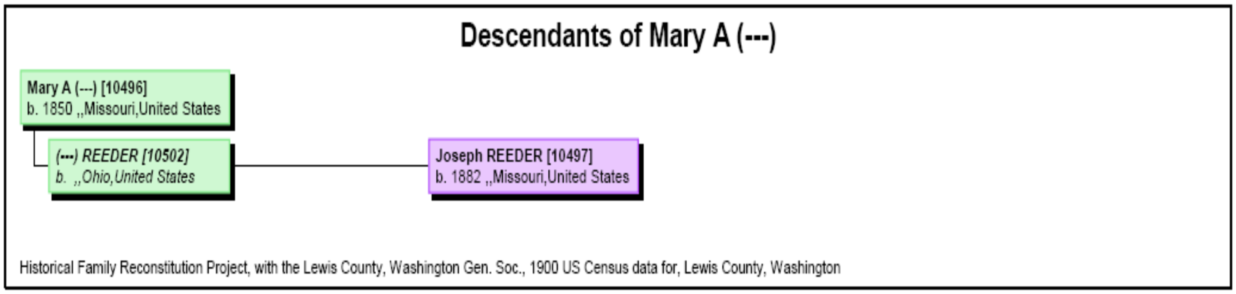


Figure 12 - US Census – wife not in the household, but her parents were, family linkage was preserved by creating a derived record for the wife.



Historical Family Reconstitution Project with, the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 13 - US Census – servant family, this household included a widow and her son, plus a servant family consisting of a father, mother, and two children. Both are captured correctly. The widow's family is displayed first (with a derived husband) and then the servant's family second.


Descendants of Mary A (---)
Mary A (---) [10496]
b. 1850 ,,Missouri,United States
(---) REEDER [10502]
b. ,,Ohio,United States
Joseph REEDER [10497]
b. 1882 ,,Missouri,United States
Historical Family Reconstitution Project, with the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

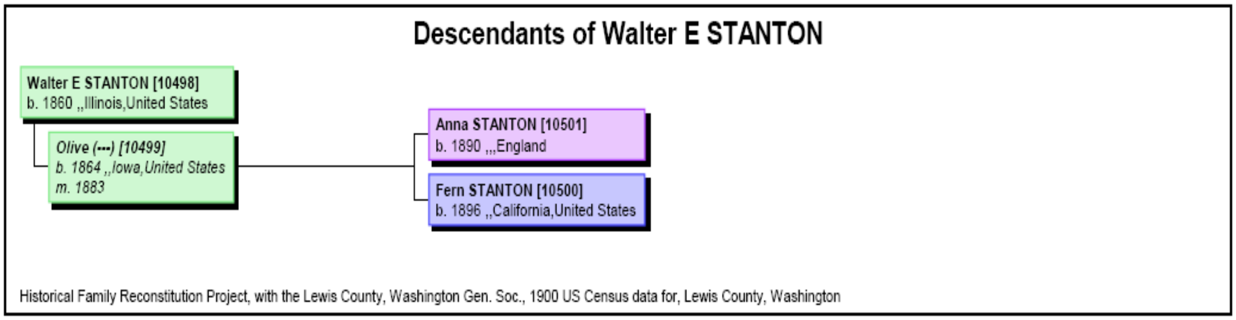

Descendants of Walter E STANTON
Walter E STANTON [10498]
b. 1860 ,,Illinois,United States
Olive (---) [10499]
b. 1864 ,,Iowa,United States
m. 1883
Anna STANTON [10501]
b. 1890 ,,,England
Fern STANTON [10500]
b. 1896 ,,California,United States
Historical Family Reconstitution Project, with the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 14 - US Census – hidden Family, male head of household with his mother present, but her surname is different than his, this suggests she has remarried since he was born, both are captured correctly. Ellsworth B. FOOTE was the head of this household, his sister also resides with him and she is correctly linked with a derived record for her husband based on her married name.

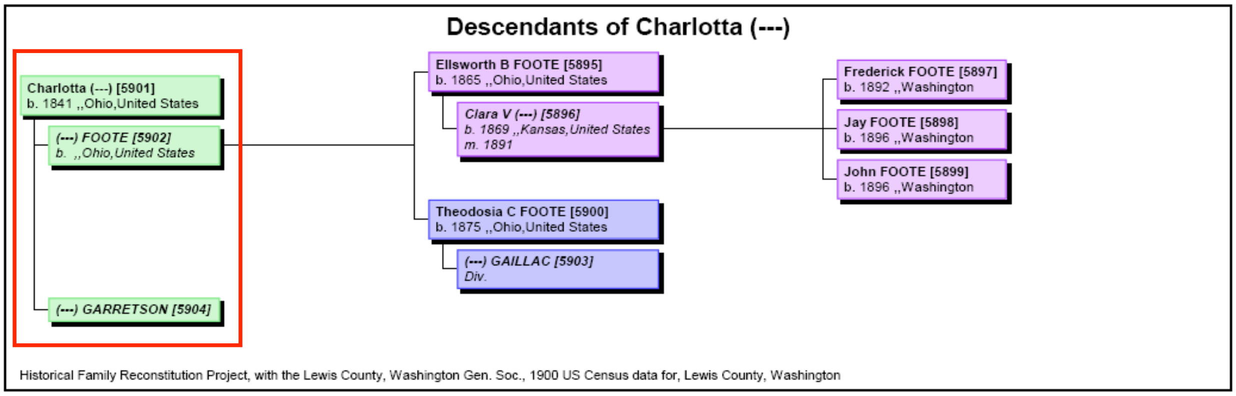
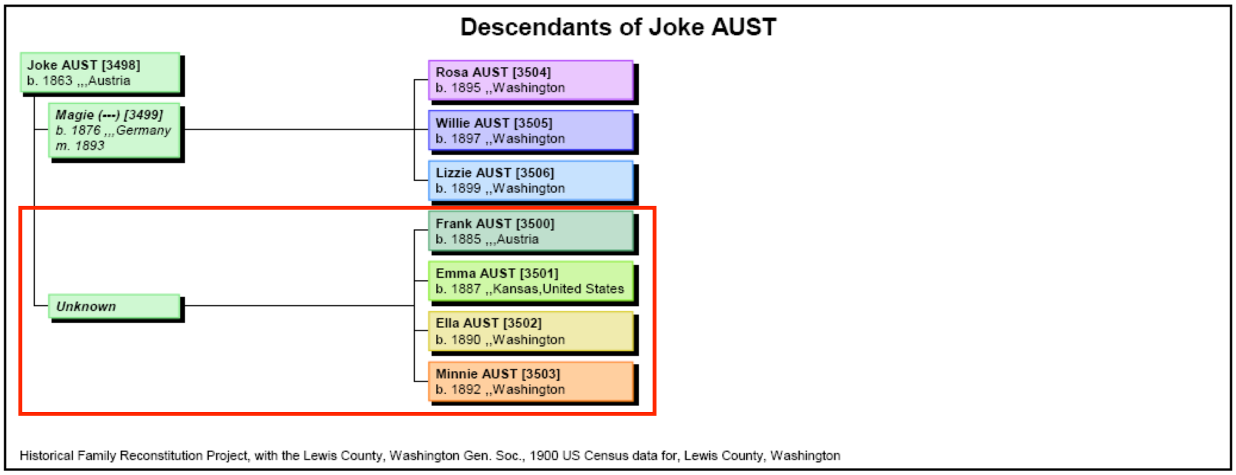


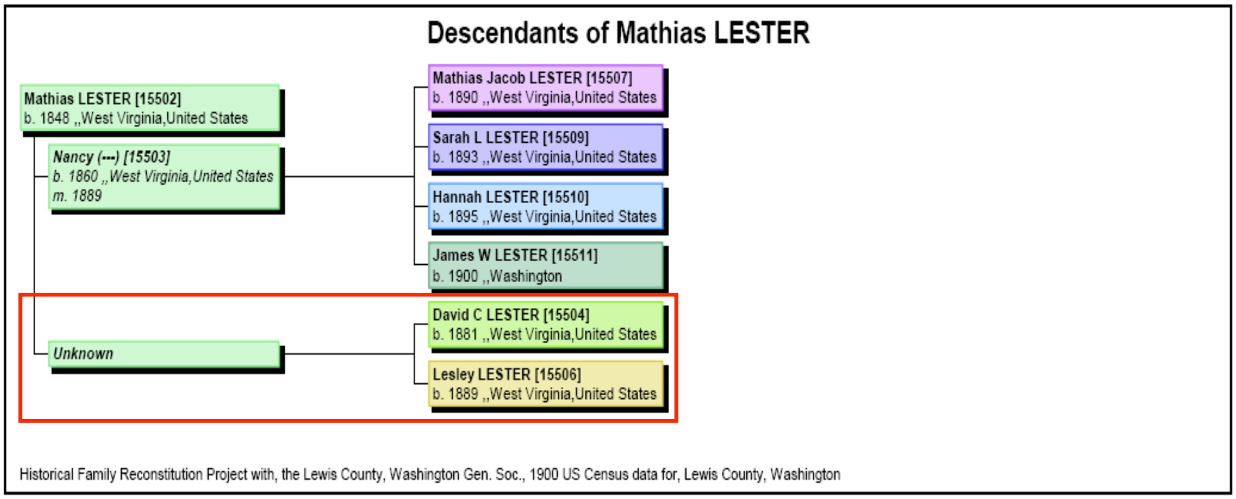
Figure 15 - US Census – hidden family, based on the marriage year in the census the husband’s children by a previous wife are in the household, by creating a derived record for the first wife both families are correctly linked. Joke AUST is the head of the household, his family from his first marriage is outlined.



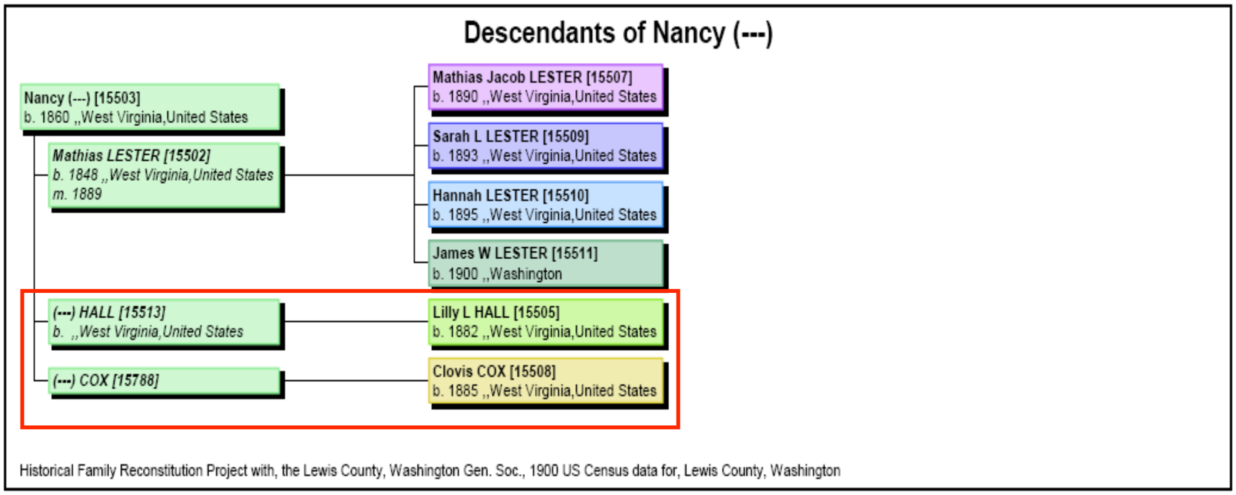
Descendants of Joke AUST

Historical Family Reconstitution Project, with the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 16 - US Census – hidden family (yours, mine, and ours). Based on the marriage year in this census this household includes Mathias LESTER, his wife Nancy, and their 4 children, plus 2 children by Mathias' previous marriage, and 2 children (one each) from Nancy's 2 previous marriages, based on their surnames. All are correctly linked. The children that are outlined are from the previous marriages.

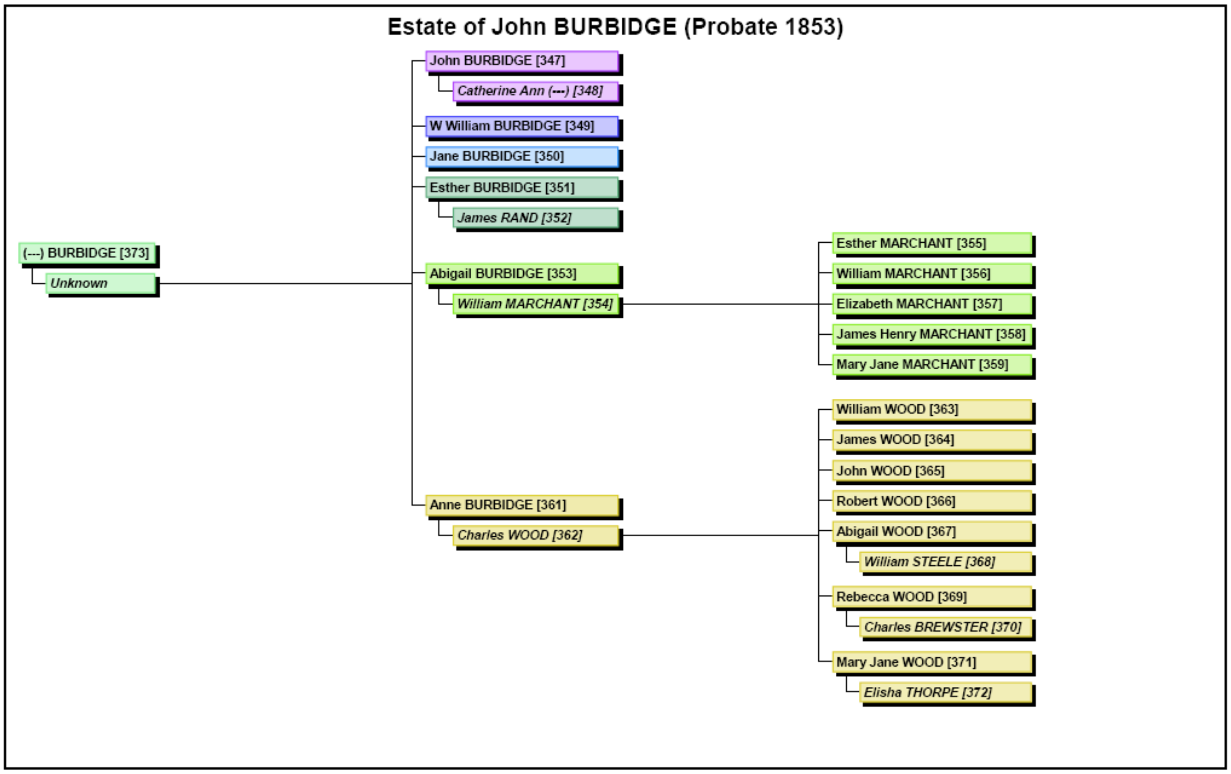

Descendants of Mathias LESTER
Mathias LESTER [15502]
b. 1848 ,,West Virginia,United States
Nancy (---) [15503]
b. 1860 ,,West Virginia,United States
m. 1889
Mathias Jacob LESTER [15507]
b. 1890 ,,West Virginia,United States
Sarah L LESTER [15509]
b. 1893 ,,West Virginia,United States
Hannah LESTER [15510]
b. 1895 ,,West Virginia,United States
James W LESTER [15511]
b. 1900 ,,Washington
Unknown
David C LESTER [15504]
b. 1881 ,,West Virginia,United States
Lesley LESTER [15506]
b. 1889 ,,West Virginia,United States

Historical Family Reconstitution Project with, the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington


Descendants of Nancy (---)
Nancy (---) [15503]
b. 1860 ,,West Virginia,United States
Mathias LESTER [15502]
b. 1848 ,,West Virginia,United States
m. 1889
Mathias Jacob LESTER [15507]
b. 1890 ,,West Virginia,United States
Sarah L LESTER [15509]
b. 1893 ,,West Virginia,United States
Hannah LESTER [15510]
b. 1895 ,,West Virginia,United States
James W LESTER [15511]
b. 1900 ,,Washington
(---) HALL [15513]
b. ,,West Virginia,United States
Lilly L HALL [15505]
b. 1882 ,,West Virginia,United States
(---) COX [15788]
Clovis COX [15508]
b. 1885 ,,West Virginia,United States

Historical Family Reconstitution Project with, the Lewis County, Washington Gen. Soc., 1900 US Census data for, Lewis County, Washington

Figure 17 - Probate Data – this same coding method has also been applied successfully to other data types, in this case probate records from Kings County, Nova Scotia. The resulting family views are exciting.



From these examples it is easy to see that capturing data sources in a lineage linked database greatly increases the availability of information and thereby the ability to link (or match) the data with other records, especially if they are in the same format.

As with any extraction effort there are potentials for errors and certain costs. Things like:

- Some census records do not have stated relationships
  - In some cases these can be added to the extracted data to allow the enhanced family linkage that is desired.
- Coding/Pointer errors can cause inaccurate family linkage
  - Our CensusToGed data converter tests for many of these errors and identifies most of them for us so we can correct them before creating the final database
- Not all families follow local name customs or patterns (like patronymics)
  - An analysis of your data set will usually reveal these exceptions
  - If they are minor, find, adjust, and allow for the exceptions
  - If they are major, we turn off that feature in the CensusToGed converter
- It takes time to populate the relationship and position pointer data fields
  - In some cases these tasks can be done accurately with automated processes
  - In other cases they will need to be done manually and you will need to decide if the added benefit is worth the additional data preparation cost

Considering the significant benefits gained from capturing the census (and other) data in a lineage linked format (families, pedigrees, and descendants), we feel it is worth the extra effort it takes to get there.

Contact Information:
David S. Barss, AG®
Project Manager
Historical Family Reconstitution
FamilySearch

E-Mail: barssds@familysearch.org
Phone: 801-240-1357