Extensible Lineage Manager
by Kevin Bottoms and Dan Olsen

Introduction

Today's genealogy programs do not provide many of the needed services for genealogists, such as searching, filling in values of search results, customizable views, adding additional information, geography help, additional file formats, publication of results, handling duplicates, etc. Some have tried to address this by creating separate programs. Genealogists are then faced with multiple programs with different user interfaces that they must learn in order to interact with each. Also there usually is not a quick and easy way to communicate information between the programs which may leave the genealogist manually entering information into each program. The alternative to typing is using the GEDCOM file format to export, import, and merge information into another program.

Our solution is to create the Extensible Lineage Manager (ELM) which is an extendable genealogy program that supports the research process. This will allow services to be easily added into the program making it possible to use just one program to accomplish genealogists goals removing the burden of communicating information between programs. Our focus on providing extension points will be in the area of supporting the ability to do research and record that research into the program.

Support for the Research Process

To support this research process we propose providing the following: estimators, suggestors, research todo list, research results, evaluator/rankers, and recorders/mergers. We will discuss each of these in turn showing their importance to the Research Process.

In order to find information it is helpful if we can narrow where and when to look for it. Estimators use known information to get estimates of dates and places of where and when events might have happened and provide alternate spellings of surnames for cases where the surname may have been altered. This reduces the need to look for missing information in all possible dates and places and provides findings where surnames have not been consistent such as in the case of immigration.
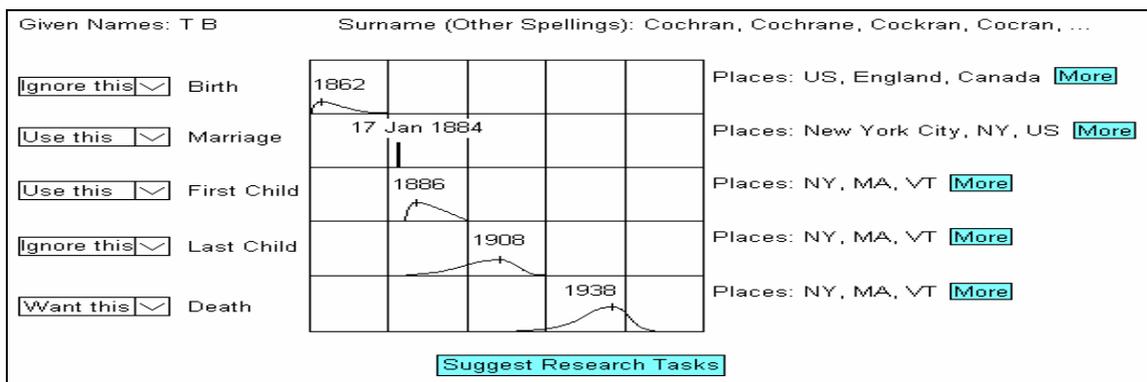
View of estimated information

There are a great number of sources of information for genealogists for specific places and dates. It is needful to know what sources apply to the given and estimated

information. Suggestors provide suggestions of research tasks that lead the user to those sources that are most likely to provide more information from what is given and estimated. A suggestor might suggest doing a search on FamilySearch for an individual or suggesting where to get a copy of an individual's will.

As the user is given suggested research tasks, he needs a way of knowing what tasks have been done, need to be done, and which he has decided not to do. Also the user needs to know how much work will be required on his part in order to do each task. Each of these tasks fall into one of three categories: Automated, Manual, and User Entered. The program will perform Automated Tasks, guide the user through Manual Tasks, and allow the user to create User Entered Tasks. A Research Todo List will provide the user with the ability to see the state of each task.

When tasks are completed the information obtained and the source need to be stored. Research Results are objects that contain what information was found and the source or task that provided the information.

There may be a great number of Research Results, especially from automated sources. The user needs a way to see the most relevant results first to reduce the need to look at all of the results. An evaluator/ranker evaluates the relevance of each Research Result and then ranks the Research Results accordingly. The evaluator/ranker could base its ranking on information gain, matching information, and probability estimates.

When the user decides to incorporate results, it is important to remember the source of information. Recorders/Mergers provide the ability to save results and incorporate them into the model. When a Recorder/Merger incorporates a result into the model, it also saves the source or task that the information came from.

```
Current Information                          Result Information
    Name: T B Cochran               ☐ [Name: T Cochran]
       Birth Date:                  ☒ [Birth Date: 28 Mar 1908]
       Birth Place:                    Birth Place:
       Death Date: Abt 1989         ☒ [Death Date: 17 Nov 1989]
       Death Place: Bethany, Oklahoma   Death Place: Bethany, Oklahoma

    Spouse: Betty Stanley
       Marriage Date:
       Marriage Place:
    Children:
       Teresa Bottoms B: 17 Nov 1953 P: Oklahoma   [Merge]        [Save]
```

Merge/Record View

Extensibility

ELM will be built on a plugin architecture to provide extensibility. This makes it easy to add onto the program by making plugins that provide extension points for more places to extend and/or extensions to the existing extension points. We will support the research process by creating extension points for estimators, suggestors, research results, evaluators/rankers, etc. We will also provide extension points for views, menu items, and multiple file formats.

Conclusion

In order for genealogists to get their research, recording of information, and other services in one program, they need an extendable program that supports the research process. With the extendable program the research process can be created and added onto (in the form of more estimators, suggestors, etc.) and the program can be added upon to provide the services that they need to accomplish their goals.