Collaborative Genealogy

Aaron Reed
Northface University
South Jordan, Utah, USA
Email (aaron.reed@northface.edu)

There are a large number of tools available to the modern genealogist which provide easy access to data, filtering, and other helpful information needed to get a good start on a genealogical foundation. However, as the size of the data grows and the number of generations identified increases, the current genealogy tools show some glaring weaknesses. These include, but are not limited to, collaboration, storing and tracking customized data, and presenting that data in an organized fashion.

Many of the tools out there are either out of date with regard to technology or don't provide the genealogist with the ability to capture and manipulate the data in a way which he or she requires. Many of the existing genealogy tools existing today suffer from both of those issues. The genealogy community needs a solution which will allow the user to represent any possible variation of data while doing so utilizing the technical standards of today and tomorrow.

In addition to this dilemma, genealogists have traditionally been somewhat isolated in both data gathering and data manipulation. As data sizes grow, however, collaboration among genealogists either within a family or more extensively becomes critical. The farther back a genealogist traces his or her ancestral tree, the amount of data and the availability of that data decreases. It becomes essential, especially when dealing with earlier generations, that a genealogist have access to easily accessible collaborative methods with which to communicate with other genealogists who may have information concerning their ancestors. While there are many tools which focus on allowing you to search and find data, not a lot of attention has been given to providing the genealogist with a way to easily manage collaborative groups as well as to easily work with those groups in hopes of uncovering more genealogical details and information together.

Finally, genealogy data quickly becomes vast and expansive even for the smallest of family trees. Add to that the complications of different family units as well as special event tracking (such as temple ordinances, etc.) and any other data which is pertinent to a particular user and you have massive amounts of data which is easily lost and/or misused. The ideal solution would provide a means to display and represent such data in ways which are both meaningful and helpful to users wishing to view and/or manipulate the data.

The Genos Project

Northface University, located in South Jordan, Utah, is a unique university dedicated to educating the most sought after software developers and informatics professionals. Northface University focuses on project based learning in which a student learns key development tactics and practices via project courses while working on internal development projects. The internal development project infrastructure focuses on the concept of people, places and events (PPE). Utilizing PPE, Northface University is able to provide its students with a large variety of development projects encompassing a number of different arenas.

One such project is known as Genos (from the Greek word, meaning "kin"). Genos is currently being developed by students at Northface University with the goal of addressing many of the issues facing the modern genealogist, including those listed previously.

The core concept of Genos lies within the foundation of the PPE model. That is, that anything related to the study of genealogy can be represented within the scope of people, places and events. The first step is to encapsulate that data and create a structure with which the data can be stored and manipulated. The resulting specification which was internally developed at Northface University to provide this functionality is PPEXML.

PPEXML

PPEXML is an XML specification which allows complete data flexibility in handling people, places and events. PPEXML allows complex data extensions for each of those types as well by handling custom data and intensive notes data and multimedia attachments to any data type. This allows users to be able to customize their data in ways previously not possible. A virtually unlimited amount of data customization allows the genealogist to keep track of data sources, observations, or other relevant information. Multimedia can be attached to any data type which allows for storage of pictures, videos, sound clips, or other media types to help store valuable memories to different people, places, or events in one's data.

One of the growing concerns among the modern genealogical community pertains to the difficulty in representing relationships in a completely dynamic model. PPEXML addresses this by encapsulating relationships in a manner which is entirely customizable and definable between any data type and another. This allows genealogists to represent data efficiently and to accurately depict any possible relationship in the modern family or elsewhere.

Finally, PPEXML allows grouping of data in any form or fashion as well. This allows the genealogist to be able to group records and individuals into limitless collections of people in order to represent data in units far more reaching than the typical family group. This becomes especially important when dealing with other users worldwide. Imagine being able to group a World War II veteran in your lineage with his or her squadron and then be able to link to those individuals and their posterity. Then imagine being able to find the descendants of those other individuals in the squadron and being given the ability to collaborate and communicate with them instantly. Couple that ability with the benefits of the dynamic relationships described previously and the genealogist could even identify individuals in the squadron with their specific roles and relationships with one another both on a personal as well as a military level.

As you can see, PPEXML offers a foundation of data encapsulation and data flexibility never before offered in a genealogical tool. Using XML as it's backbone for data representation, Genos also has been established on the latest technological tools and specifications which will make extremely flexible as future technologies evolve. Using PPEXML as it's foundation, Genos is afforded flexibility and customization in data representation never before available to the genealogy community.

Genealogical Collaboration

As data sizes grow and as genealogists trace further back their ancestral tree, the difficulty in locating and confirming new ancestors becomes more and more difficult. One of the best and yet most underused tools available to genealogists comes in the form of other genealogists and the data which they have amassed. Two of the key reasons why such a great tool has been underused are: 1) the inability to locate genealogists with similar heritage, and 2) the difficulty in communicating with other genealogists in order to facilitate data sharing.

Genos addresses the first issue by providing a flexible yet powerful collaborative partner search. Essentially, this tool allows the genealogist to identify individuals in his or her ancestral tree and then search other genealogical data in hopes of finding potential matches in the trees. The search results are represented with a match percentage (representing the likelihood that the person found is indeed a matching record) based on a number of different factors which the genealogist is able to manipulate. This tool will be an essential piece of ongoing genealogy and will allow genealogists from across the globe to reach one another and open portals of communication which will aid in family research worldwide.

Genos takes things one step further in providing the genealogist with several tools to provide better and more efficient communication with collaboration partners once they have been located. Genealogical based solutions via email, instant messaging, generic

message board forums, individual record forums, instant collaborative chat, and others provide users with unprecedented access to a large number of resources which will instantly tie them to their collaboration partners regardless of their physical locations.

Data Representation

Along with the benefits of providing unmatched flexibility in a data model comes the drawbacks of the technical difficulties in representing such data. As the dynamics of the data grow and the boundaries of the data become more and more far reaching, the difficulty in understanding and harnessing the data increases. Genos provides a number of different data views which allow the user to depict his or her data on screen or on paper in ways which accurately and clearly illustrate all the different data types and relationships which Genos allows the user to create.

Genos Synopsis

As an internal development project, Genos provides students with an exceptional learning experience. Students work in teams and under the influence of team leads with a professional reporting structure. This provides the student a great working experience and allows them to get accustomed to working in a professional environment. Genos is written using the latest Microsoft technologies (.NET and SQL Server 2000), Java technologies (IBM Websphere) and other current standards (XML, XML Web Services, XSLT, etc) which also gives the student great exposure to current hot technologies while working in a professional setting. In serving as a student project, Genos accomplishes a twofold mission: 1) Genos is a means by which Northface University can allow it's students to grow and develop and therefore reach it's goal of educating the most sought after software developers and informatics professionals; 2) Genos addresses a number of the major dilemmas facing the modern genealogist as described in this paper. The future of the genealogical community will be greatly enhanced as Genos continues to evolve.