ProgenyLink.com
Software Function Layers

Reengineering the
Fragmented Genealogy Industry

Turning the Internet into a
Single Integrated Genealogy Database

Reengineering the Fragmented Genealogy Industry

The genealogy industry is currently extremely fragmented in its operations, and it would greatly benefit from adding more cooperation and synchronization to the entire process. This fragmented condition is evident in the normal practice of researchers working alone. These millions of genealogists tend to work alone presumably because the pedigree-sequence method of research makes it extremely difficult to cooperate and coordinate with anyone else. In most cases, any two researchers may share no surname lines at all, or may share only one or two, and the timing of their interest on such surname lines rarely coincides.

This fragmented condition of the industry is also evident in the way that researchers display their data for others to access. This practice of displaying data in public is an idealistic and ambitious effort to cooperate with others on a purely volunteer basis, but it has its problems. On March 17, 2012, cyndislist.com showed links to 309,968 separate websites, indexed in 188 categories. I have not investigated whether these are all free sites or if some of them require a subscription or other payment. But I know from discussions at genealogy conferences, that there would be many more websites containing useful information if there were an easy way for researchers to display and market their data online. For example, one gentleman photographed and digitized every document in a county courthouse. He would like to put that data online and to do further digitizing of public records, but he needs a way to market that effort to pay for his expenses in this idealistic effort. If there were some mechanism which made it easy for him to market that data, then he would be encouraged to offer even more data for the benefit of genealogy hobbyists. Apparently, the Internet companies with large collections of public records are not willing to host and market data for others with small collections.

It seems like it would be ideal to offer a service which allowed some of these hundreds of thousands of websites to place their data in a common location, in a common format, with a common indexing method. For example, that might allow 200,000 of the 300,000 websites listed on Cindislist.com to be integrated into a single central database. As you can imagine, an integrated version of all these websites should greatly ease the task of genealogy hobbyists who would like to take advantage of these public displays of family research work. Google probably gives reasonably good access to each of these websites individually, and offers an "integrated view," in the sense that all similar data is listed together. However, there are many more steps which could be taken to minimize duplication, etc., if there were a place to combine this data properly.

Stepping through the layers of a complete software solution

The ProgenyLink.com system is intended to solve these problems of industry fragmentation, along with a host of other major problems in the genealogy industry. Below is a list of the system elements which are gradually being stitched together to make the complete system. We hope that many genealogists will take advantage of the work we have done so far, and become supporting members of this project. That will allow us to continue until we have finished this grand vision in its entirety.

Main System Elements

The completed version of the ProgenyLink.com system will contain the following elements:

Completed?	System Element
Yes	1. PAF replacement – genealogy data entry and image storage online.
Yes	2. Add centralized cooperation mechanisms (to make PAF into "SuperPAF").
No	3. Bulk input and output system – Also, isolate data where necessary for quality reasons or separate administration.
No	4. Finance – sell and trade names and source data – provides fairness and income incentives to finish the whole nation.
No	5. Incorporate off-site data into the indexing and retrieval system.
No	6. Charge for off-site data, in some instances.
No	7. Semi-automatic sourcing and descendency compilation.

The System Expansion and Generalization Process

Semi-Automatic Sourcing and Descendency Compilation

Charge for Off-site Data

Incorporate Off-site Data

Financial System - Fairness & Incentives

Bulk Data In and Out

SuperPAF-Cooperate

PAF