

Sequential Process View of Invention – Reengineering the Entire Genealogy Industry

This requires multiple technology breakthroughs, all at once

Theoretical efficiency improvements of up to 2000 times are possible, as in any typical mass production/industrialization process, in contrast to the inefficiency of typical cottage industry methods. Actual improvement in the range of 30 to 100 times should be easy to achieve. (Mass production techniques have never been applied before to the crucial genealogy process of name assembly. Adam Smith, *Wealth of Nations*, put 4800 as the top industrial improvement multiplier observed.)

Register/setup	Enter Data	Assign numbers	Store names in descendent-sequence structures	Data owners control access and updates to data.	Continual data improvement process to reach required quality	Connect Surname Groups through Women	Data Quality Filter/Barrier	Pay-Per-View Data	Sale of Names/Final State of Database
<ul style="list-style-type: none"> ● (1) Register (identify) user, register surname and associated ancient ancestor, allocate workspace, pay membership dues. ● (2) Establish multiple workspaces for ambitious users or workgroups, each workspace used to store a different surname group. 	<ul style="list-style-type: none"> ● (4) Direct entry from prior manual research. ● (5) Bulk GEDCOM input ● (6) Employ specialized semi-automated assembly of all index entries, names, and related documents of potential interest for single surname. (Only possible using descendent-sequence system). ● (7) Use "process of elimination" separate subsidiary database to show which public record images have been used in the main database. ● (8) Broker and coordinate outside research. 	<ul style="list-style-type: none"> ● (9) Assign worldwide unique number for all dead and living, as data is entered. Turn the Internet into one integrated genealogy database with a unique ID for each possible person. See 4-level number, which identifies data owner, and keeps descendency groups separately addressable throughout the database. 	<ul style="list-style-type: none"> ● (10) Store names in descendent-sequence surname groups. – avoids nearly all duplication (data owners enter data for only one surname -- "Descendents of...") and notify other participants of each user's research intentions – allows industrial-strength cooperation across surname lines. 	<ul style="list-style-type: none"> ● (11) Specialized workgroup networking features for genealogists including multi-level access rights. ● (12) Provide special provisional update methods. ● (13) provide special temporary workspace or "shadow database" transition and transformation processing space to support numerous special transactions. ● (14) Data ownership is recorded at the name and data element level. (optional) ● (15) Provide "Everyone can update" feature for "community data" projects. 	<ul style="list-style-type: none"> ● (16) Improve data in normal ways. ● (17) Semi-automatic mechanism to find source records to link to previously assembled names, using "screen scraping" and other techniques ● (18) Unique document and image upload process by individuals. ● (19) Use public catalogs as input to source-identifying entries. ● (20) Assess current data quality levels as needed. ● (21) Special transactions, part of improvement process. ● (22) Link names to source records which are uploaded to GenReg. ● (23) Link names to source records on major sites with stable URLs. ● (24) Link names to various websites cross-indexed to GenReg, containing videos and other voluminous data. 	<ul style="list-style-type: none"> ● (25) Provide internal email system ● (26) Process for connecting surname groups must use high-quality data to avoid confusion and wasted effort from unstable data. – users receive 1000-to-1 return on data entered and connected as other users provide 10 generations of data on 1024 surname lines. 	<ul style="list-style-type: none"> ● (27) Categorize data by quality, and search and list results by quality category: Up to six categories of quality. High, medium, low individual quality, plus size of network of related names. 	<ul style="list-style-type: none"> ● (28) Confirm data has reached Pay-per-view quality levels, the final step in the quality improvement and recognition process. Diligent users receive: – 1000-to-1 data return on data entered. – Eligible to receive royalties on marketed data. 	<ul style="list-style-type: none"> ● (29) Sale of finished data. Internal financial system tracks all user-related transactions. ● (30) Royalties go back to data supplier – adds big incentives to finish whole nations by filling in all the data gaps. Users pay for data they download, and receive payments for data others buy. System pays net royalties periodically. ● (31) Collect and remit payments for "on consignment" data on other sites (see item 23). ● (32) Record stripping -- The final state, an Historical "Facebook" for all historical people. ● (33) History-based social networking. ● (34) A more accurate method for indexing source records.

● (3) Users receive royalties to offset membership dues and may receive net positive cash flow.

(30, 31)
Royalties go back to data supplier (profit-sharing)

(30, 31)

Extended Explanations

(6) Workstation software and hardware uses sophisticated "screen scraping" and other techniques to find and assemble all index entries and names from numerous online "raw data" databases and related source documents for a specific surname.

This process greatly accelerates the manual review and assembly of family structures at local PC document retrieval speeds which may be up to 400 times faster than unpredictable Internet speeds. Users can view dozens of documents simultaneously, while making comparisons among them, where useful.

This process can only be done using descendent-sequence (single-surname) method, and puts results in database in descendent sequence. (There is no practical way to do this with pedigree-sequence research, since new, usually unknown, surnames are introduced at every step backwards in time. For example, there are 1024 surnames needed at 10 generations back in time.)

--See conceptually related process under "continual data improvement," item (16), where source records are semi-automatically located for previously assembled name structures.

(9) Using a unique "tree-level" number, also allows entry of lists of names which are not connected into family groups. These names can later be assembled into family groups using the "descendency-level" number. This might include such things as the lists of Russian prisoners sent to death camps, where Russian genealogists have preserved those names, but have not yet included them in pedigree structures.

(11) Specialized workgroup networking features for genealogists including multi-level access rights. Access rights of View, Provisional Update, and Immediate Update are granted to family and friends and provide numerous workgroup networking features, specifically for genealogists. Same-surname cousins (who all have a common ancestor) should be the first group to invite to assist.

(12) Provide special provisional update methods to assure shared responsibility and control to achieve high quality. Includes option to review of all updates after-the-fact, by date and operator.

(13) provide special "shadow database" transition and transformation processing space to support numerous special transactions, including various provisional update transactions, as in item 12 above, item 15 below, item 21 below, etc.

(14) Data ownership and responsibility is recorded at the name level and at the individual data element level.

(15) Unique, carefully controlled "Everyone can update" feature for international "community data" projects.

(17) Semi-automatic and manual location of sources – link names to online sources or to sources uploaded to Genealogy Registry. This is similar to process used in item (6).

(18) Participants can upload personally acquired documents. (This should end most institutional, contractual, archival, media, and structural barriers to location and use of the world's genealogy source documents.) Scanners, cameras, downloads from various scattered sites, etc., may all be inputs to this process.

(19) Use public catalogs as input to source-identifying entries.

(20) Assess current data quality levels as needed

(21) Special transactions, part of improvement process. Collect data fragments from throughout database into one work space. Connect hierarchical workspaces with each other.

(26) Women usually appear twice in the database. A woman in the role of a daughter is connected to that same woman in the role of a wife in a separate surname group using "same person" links. This ties all portions of the database together so that all possible pedigrees can be read out at the end of the database construction process. Databases are constructed in descendent sequence, simply because that process can be hundreds of times faster than the traditional pedigree-sequence methods.

Some general benefits and consequences.

- Cooperating across surname lines is the most powerful benefit of all. (26)
- (35) Engaging genealogists worldwide will maximize the cooperation and achievable gains. General enthusiasm from expecting quick completion.
- Turn the Internet into one integrated genealogy database. (9)
- Choose which data should be visible to public searches. Avoid most confusion. (27)
- Semi-auto assembly of family structures. (6)
- (36) Encourage organizing family organizations, usually of same-surname cousins.
- Centralizing the indexing and marketing of data on 3rd-party websites. (24)
- "Records stripping" -- a nationwide correlation of all historical records for individuals, creating an "individual level" national history. (32)
- Acquiring all the world's genealogically significant source records images through individuals. Improve legal access to more genealogical data. (18)
- Researching in descendent sequence is very efficient. Avoids duplication of research and duplication of names in database. (4)
- (37) Reclaiming the 20% to 25% of people lost to pedigree-sequence research.
- End of email. Users rarely need email any more, and if they do, it is internal. (25)
- (38) Database entries can be reverified hundreds of times to improve accuracy.
- Financial subsystem -- gives revenue, increases fairness, adds new incentives to finish nation. (30, 31)
- (39) Achieve industry business integration for another layer of efficiency.
- Workstation for semi-automatic finding of source documentation. (17)
- (40) End most "brick walls" in genealogy research. No ad hoc "reverse gen."
- (41) Solve problems for LDS Church, and they put more money into the project.
- (42) Solves all industry technical problems at one time: duplication, cooperation, integration, fairness, quality, uniqueness -- end weak incremental improvements.

(3)