Interactive Genealogy Explorer: Visualization of Migration of Ancestors and Relatives

Daniel Shakespear

Independent Programmer
Munich, Germany
E-mail: daniel.shakespear@gmail.com

Abstract

An interactive, three-dimensional graphic interface was conceived and developed, to illustrate and enable exploration of events recorded in a genealogy. The resulting visualization provides a colorful and intuitive view of migrations, as well as plotting the biographical data for individuals in the genealogy. Users can quickly highlight the connections between people, get a perspective on which persons lived in similar eras, or trace the life events of any person in the genealogy.

The events of individuals' lives are plotted as colored lines above the backdrop of a globe (Google Earth), tracking the life and travels of the person. The date of each event determines the altitude of the line at the event location, and the lines are colored corresponding to the date of birth of each individual. The user can manipulate the view manually, or cycle through the events in the individual's life. One lifeline or up to hundreds can be shown (for example, chosen by date, name, ancestors, descendants, connection between two individuals, etc.), to give a sense of the connections between people, and provide a visual overview of migrations and movements.

Keywords: visualizing biographical data, automatic biography generation, visualizing genealogies, Google Earth, Google SketchUp

1. Background

Origin has long been important to individuals and cultures across the world. Whether it be the lineages described in the Bible, the peerage system and countless other mechanisms of hereditary power transfer in Europe, or the massive genealogy of Confucius (first published in 1054 AD, and most recently updated in 2009, spanning 2500 years and more than two million members), pedigree has helped establish identity and authority in societies through much of recorded history.

Even in the United States, where birthright has arguably played a lesser role in determining status, genealogy has a long and steady history of interest. Individuals' fascination with the age-old question of provenance has flourished in the internet era. In 2005, a survey indicated that 73 percent of respondents were interested in researching their family trees, an increase of 13 percentage points in only five years (Weil, 2013). Between the online availability of mass quantities of genealogical data, and the development of user-friendly software (a popular rating website counted 906 genealogy application as of late June 2015)[1], professionals and laypersons have unprecedented possibilities to explore ancestry and kinship.

Over the years, genealogies have been visualized in various ways, each with advantages and disadvantages as a function of the data set and the intended interpretation. The classic family tree and (Figure 1) is well-suited to show the heritage of an individual. Typically, parents and direct ancestors are the primary focus, not siblings. A radial chart can also be used, with one generation plotted per ring (Figure 2). This provides a more-compact view which is still focused on the central individual.
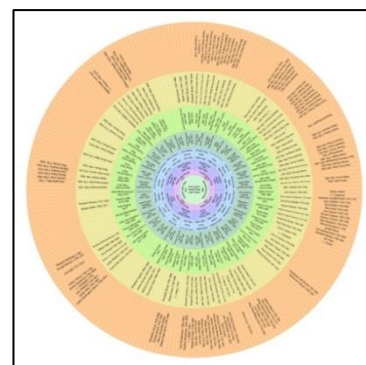

Figure 1: Traditional family tree[2]

Figure 2: Radial family tree[3]

The traditional family tree diagram can be extended with siblings (Figure 3) for a comprehensive view of a group of individuals and families. The resulting amorphous shape lacks the symmetry of the simpler, individual-focused charts, but enables a broader range of connections and relations to be portrayed. The more-complex GeneQuilt (Figure 4) displays this sort of information in a horizontal

---

1 http://www.gensoftreviews.com/index.php accessed 28 June 2015

2 http://www.scifres.com/2010/01/genealogy-visualizations accessed 27 May 2015

3 ibid.

format, and is suitable for complex family relations (multiple spouses, intermarriage, etc.). The interpretation of this chart is not, however, as immediately clear to the new observer as that of a traditional chart.
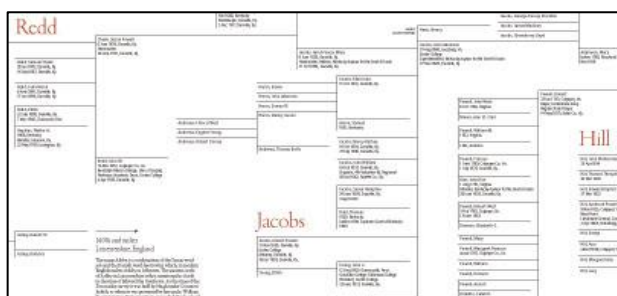

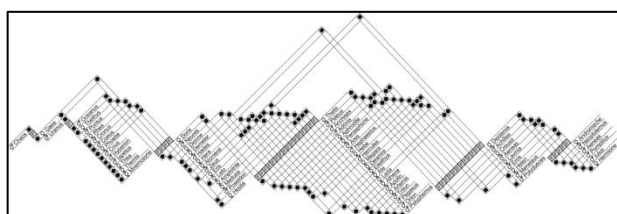Figure 3: Traditional tree with multiple parallel families[1]

Figure 4: GeneQuilt[2]

These traditional visualizations share several common traits: they are all two-dimensional, generation-based (an individual's placement is determined by their generation, not their year of birth), and limited to showing events births, unions and (when indicated) deaths.

More recently, consumer software for genealogy generation has enabled entering additional details of individuals' lives, and also geographic information about their life events. The de-facto industry file standard GEDCOM (The Church of Jesus Christ of Latter-day Saints, 1995) supports the storage, importing and exporting of this information as well, enabling users to develop and share genealogies with much richer detail than traditional charts provide. This information can be displayed by the software as a list of events, or a map of locations.

A new project developed in the last two years, RootsMapper, takes this a step further, showing migration of individuals as lines above a map (Figure 5). This is one of the few examples of a genealogical visualization which also illustrates geographical relationships.
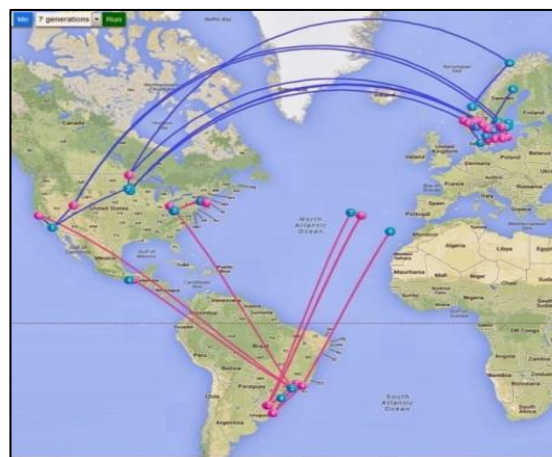

Figure 5: RootsMapper[3]

2. Concept

While inputting a genealogy of several generations of his family, the author realized that much information about the individuals was not readily visualized by the existing tools. A three-dimensional, interactive illustration of the lives and events of the people was envisioned. Ideally, the visualization would make it possible to grasp the time and location of the people and event shown, along with their connections to each other. It would serve the purpose of finding and graphing the connection between related persons. Finally, it would also enable the user to explore events beyond the normal scope of a genealogy tree, such as journeys, accomplishments, occupations and the like, and give a sense of chronological time to the data (not only generational increments).

The visualization that was developed, the Interactive Genealogy Explorer, plots events of selected individuals' lives as colored lines above the backdrop of a globe (Google Earth). The date of each event determines the altitude of the line at the event location, and the lines are colored corresponding to the date of birth of each individual. The user can move the view around manually (exactly as in Google Earth), or cycle through the events in the individual's life. For each event, a short description is shown, and the user can zoom in to see the location in detail. Various methods of selecting individuals, navigating through the genealogy, and showing paths between persons are available for interactive exploration.

1 http://infosthetics.com/archives/2006/07/largescale_genealogical_history.html accessed 27 May 2015

2 http://www.aviz.fr/geneaquilts accessed 27 May 2015

3 http://janasgenealogyandfamilyhistory.blogspot.de/2013/10/tech-tuesday-rootsmappercom-ya-this-is.html?spref=fb accessed 27 May 2015

3.   Initial Technical Implementation

To display a genealogy, the user first enters the data for the individuals using standard consumer software / freeware capable of outputting data in the GEDCOM format, such as My Heritage Family Tree Builder. By the nature of the visualization, events to be plotted require a date and location in order to be positioned. Once the genealogy is complete, the data is exported to a GEDCOM format file.

Once the GEDCOM file is available, the visualization is created in two steps, using commercial freeware, and author-developed applications. An overview of the applications and code modules is found in Figure 6.

3.1   Geometry Generation

The first step, shown in the upper box of Figure 6, is the automated reading and interpreting of the GEDCOM file, followed by the generation of the lifelines. For this task, Google SketchUp is used, with a Ruby script geometry calculator, and a web browser user interface. The user can select which individuals to plot, along with the timescale; additional options are available in the Ruby script (color schemes, direction of altitude which indicates time, etc.). Locations are converted to latitude/longitude using the geonames.org lookup service, and dates are converted into altitudes. With this information, the code steps Google SketchUp through the generation of the individuals' lifelines and accompanying text, and exporting the CAD data as COLLADA files. A camera location file is also generated, with viewpoints for each event.

3.2   Displaying the Visualization

Once the geometry is generated, the interactive visualization is available via a web browser (code and applications shown in the lower box of Figure 6). The Viewer UI module, written in Javascript and HTML, is opened in the browser (here, Internet Explorer) and it instantiates the Google Earth Plug-in to show the globe. The user then can browse through the individuals from the genealogy, and show one or up to hundreds of lifelines over the Google Earth globe. Figure 7 shows a sample view of one lifeline.

3.3   Navigating the Visualization

A variety of navigation methods are available. From any individual, the user can navigate to his or her parent, spouse, or child, and continue to explore. Alternately, multiple life lines can be selected and shown (for example, by date, name, ancestors, descendants, connection between two individuals, etc.), to give a sense of the connections between people, or to provide a visual overview of migrations and movements.

To illustrate the connection between two individuals, the path between them can be shown as a trail of lifelines. The user can move from one person's events to the next,

stopping to read events in the individuals' lives, or zooming in to see the town where an event took place. The Interactive Genealogy Explorer also generates clickable placemarks in space near the description for each event, allowing the user to click on an event and go directly to it.

4.   Development and Refinement

After developing an initial visualization in early 2012, a number of improvements were made. Initially, color was not used to indicate time, but rather was selected from a set defined in the code. Only the altitude of an event graphically indicated its timing. It was quickly clear, though, that altitude is difficult to assess by eye, especially in a spherical coordinate system such as a globe's. In order to enhance the perception of time, the color of an individual's lifeline was then selected by their date of birth (or first graphed event), using a rainbow progression (red for latest through blue for earliest individuals).

A further enhancement was the addition of white lines from the events to the surface of the globe. Analogous to the discussion on altitude, it was also initially difficult to judge location on the surface, for events plotted high in space.   The dropping lines are a very effective improvement in this regard, and also highlight the geographic proximity in the case of events which took place in different time periods.

A significant limitation to interaction was the absence of a click-event from the Google Earth Plug-In when the user clicks on a lifeline. Users are accustomed to being able to click directly on symbols and get information, navigate, etc. Unfortunately, the Google Earth Plug-in does not support this functionality for imported objects it is displaying, such as lifelines. It does, however, allow for placemarks to be shown, and reports click events from these.   To enhance the interactive nature of the visualization, placemarks were added in space at each event (when possible, with a flag of the nation in which the event occurred).   When the user clicks on the placemark, the visualization flies to its associated event. An example showing the placemarks is in Figure 8.
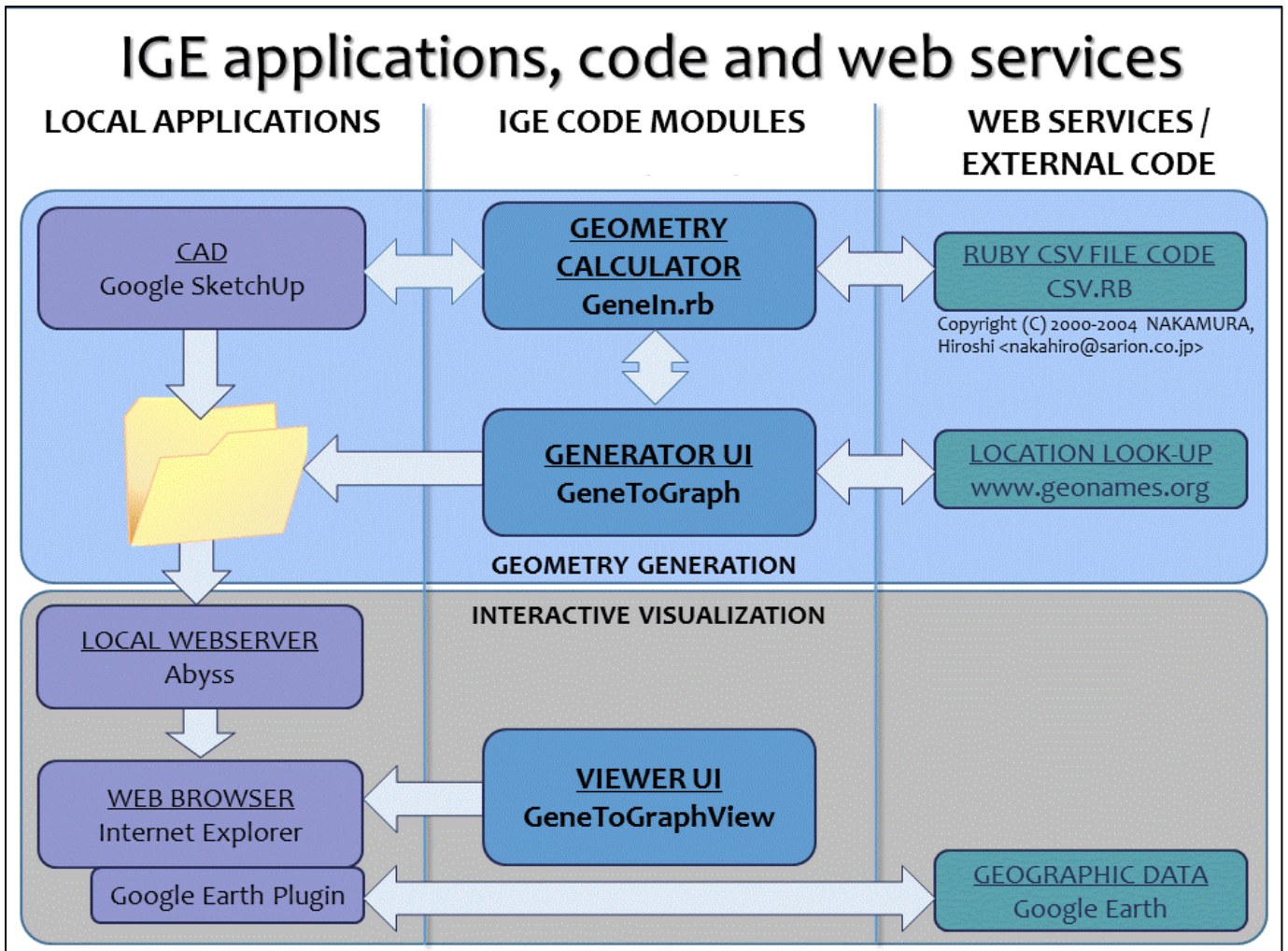
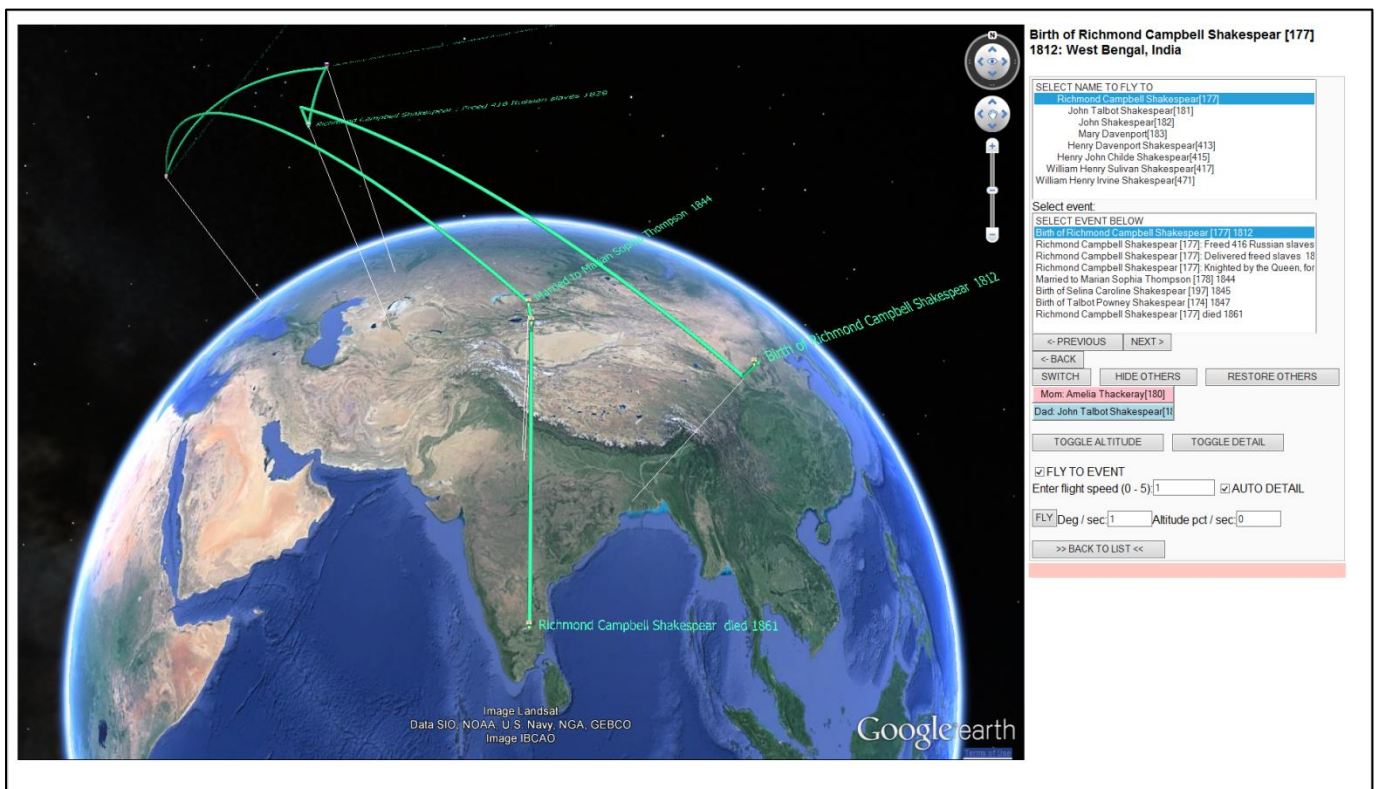Figure 6: Code modules, applications, and data flow

Figure 7: Interactive Genealogy Explorer showing the lifeline of Sir Richmond Campbell Shakespear
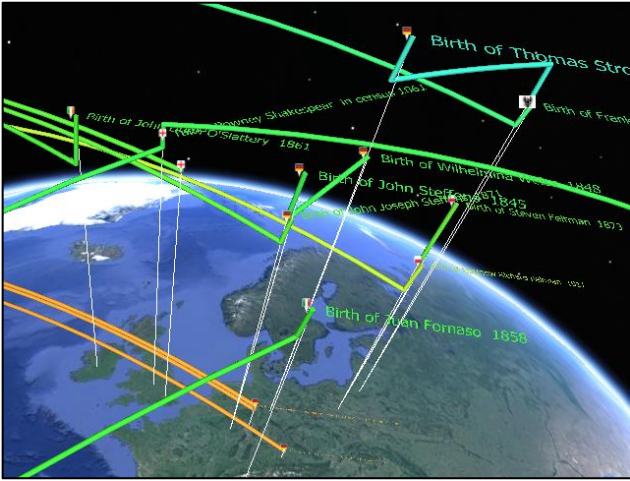
Figure 8: Clickable placemarks (flags) for events

Another major challenge was visual clutter. For example, in the case of a family with multiple children, the initial visualization was impossible to read, as the texts for each birth were plotted directly above each other. This was corrected by checking each event's distance to other plotted events. In case the event is close in time and location to an existing event, the angle of the text is incremented, resulting in a spiral appearance, and significantly better readability, as shown in Figure 9.

Figure 9: Spiraling of names in close proximity

5. Results and Comparison to Traditional Visualizations

Traditional two-dimensional visualizations (as shown at the beginning of the paper) are well-suited to quickly read lineage, as they are based on generation increments. Visualizations which show parallel individuals (not only one person's direct ancestors) can be used to explore more complex relationships (such as cousins, multiple unions, etc.). In typical consumer genealogy software, some degree of interactivity is enabled, to allow navigating through families or finding paths between individuals. Finally, traditional two-dimensional visualizations lend themselves well to printing for books or for static display.

RootsMapper, released in late 2013, provides an appealing display of basic ancestral migration. It is nicely integrated in the framework of a consumer genealogy tool (FamilySearch.org), and unlike the Interactive Genealogy Explorer, it requires no local webserver. The information displayed, however, is limited compared to the visualization developed here, the graphic itself is fundamentally two-dimensional, and the map lacks the navigational possibilities of Google Earth.

The Interactive Genealogy Explorer has been used primarily by the author (although application to other genealogy or biographical projects is welcomed). As such, evaluation has been limited to personal opinion, and feedback obtained from informal demonstrations. By its nature, the Interactive Genealogy Explorer is more difficult to compare to traditional visualizations, as it serves a different purpose. Initial observations show that it provides a richer display of information about the individuals, particularly among families with a history of migration. Its interactive nature encourages exploration, and the interface is fun and intuitive to users of Google Earth. The visualization provides the opportunity to see the places in which events took place (at least as they appear today), and make discoveries about the proximity of otherwise distant relatives.

One example of this occurred when it was discovered that the author's brother had lived in California just over 100km away from a previously-unknown third cousin, for several years. The cousin was a descendant of an individual who migrated from India to the US in the 1800s, and whose descendants had long fallen out of contact with the author's branch of the family. The visualization clearly highlighted this previously unknown connection (Figure 10). Similarly, the visualization helped the author discover that his current residence is not far from that of a distant cousin, Dorothy Shakespear, who lived recently in Italy. Figure 11 shows the entire path from the author to Dorothy Shakespear.

Figure 10: Distant cousins, related to a common ancestor in India, in close proximity.
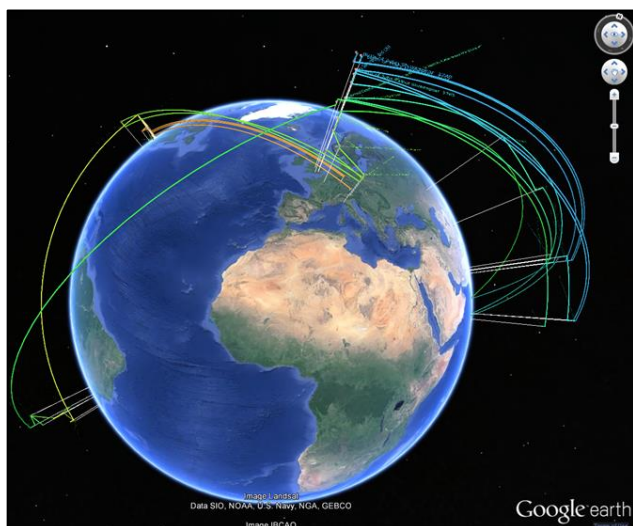
Figure 11: The author's connection to Dorothy Shakespear

6.   Limitations and Further Work

As with any visualized data, care must be taken to assess any conclusions within the limitations of the completeness of the input data. For example, if extensive data is available about one individual or one particular branch of a person's ancestry, it draws focus to their location and history, while the other unknown ancestors and their history remain invisible. This can lead to bias in the impression and interpretation of an individual's heritage.

By nature of the geographical aspect of the Interactive Genealogy Explorer, the visualization results are not as interesting for families with minimal migration, and the names and events become more difficult to distinguish in the absence of geographic displacement. Figure 12 shows this effect.
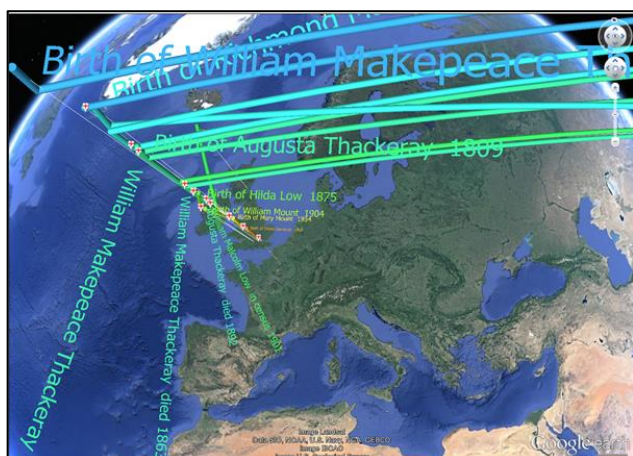

Figure 12: Connection of David Cameron to William Makepeace Thackeray

Historical changes in national borders present a challenge as well. The visualization shows events accurately positioned on the globe by latitude and longitude, but Google Earth displays nations and borders in their current

state, not accounting for different earlier political divisions (e.g. Prussia).

Technical issues include limitations in Google Earth Plug-In interactivity, and its requirement for a web server to load data. In addition, the Google Earth Plug-In is due to be deprecated in December 2015, and its replacement, if any, is currently unknown. The visualization will likely need to be significantly recoded to function after this time.

As of June 2015, further images and videos of the Interactive Genealogy Explorer are located at the project site,
https://www.behance.net/gallery/22589523/Interactive-Genealogy-Explorer

7.   Further Ideas

7.1   Concept: Scalable Translucent Globe

To make it easier to study specific times in history, a translucent, scalable globe with time-appropriate maps could be developed. The radius of the globe would correspond to the time scale of the genealogy data (altitude), and the user would size the globe to correlate to a year of interest. This would illustrate the geopolitical borders as they were, and highlight the people alive at that time (newer events and people would be visible, but dimmer inside the translucent globe). Such a feature would require a different source for globe data, as the Google Earth Plug-In is not able to render scalable dynamic objects.

7.2 Concept: Visualizing Traditional Wars

Another possible application would be to plot battles and military movements in a traditional war. This would provide a three-dimensional overview of the fundamental events and movements of troops and weaponry. Their relative numbers could be indicated by the size of the tubes, and the impact of a battle on each army would be displayed by the change in size of the tube after each battle. A translucent map which could be set to any altitude (time) with the corresponding fronts and borders shown, would give a snapshot in time at any desired phase of the war (here a flat map, rather than a globe, would likely be appropriate).

7.3 Concept: Visualizing Mass Migration

A larger-scale application could be developed to graphically show the progression of migration over time (e.g. migration of early ancestor groups out of Africa, or the migration due to the 19th century Potato Famine). In this case, the size of the shapes at a given height would correlate to population at that given time. The map in Figure 12 shows when Homo sapiens arrived in various areas over time. A three-dimensional version with time as the third axis, would allow the depiction of populations as shapes spreading out in space and time.
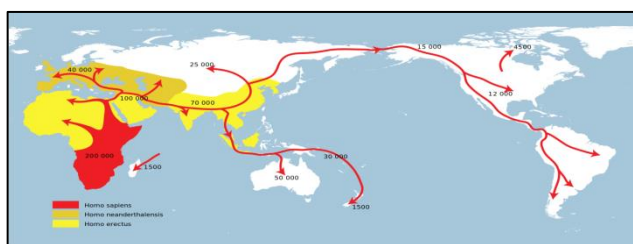
Figure 12: Migration of Homo sapiens [1]

If such shapes were generated for all three species, it would also provide an illustration of when and where they interacted. An ambitious programmer with the right data could generate dynamic sections or shapes, to visualize population sizes at a point in time.

Such an approach could also be used to track and visualize the spread of languages, diseases, ideas, or anything which permeates the planet and its people.

8.   References

The Church of Jesus Christ of Latter-day Saints. (1995). *The GEDCOM Standard, Release 5.5*. Retrieved 28 June 2015 from https://familysearch.org/developers/docs/gedcom/gedcom55.pdf

Weil, Francois. (2013). *A History of Genealogy in America*, Cambridge,MA: Harvard University Press.

---

1 http://commons.wikimedia.org/wiki/File:Spreading_homo_sapiens_la.svg accessed 27 May 2015