

Organization: Historical Society of Pennsylvania

Primary mentor: John Houser, Chief Information Officer, Historical Society of Pennsylvania

Supporting mentor: Caroline Hayden, Digital Services Manager

Project title: Enhancing Access to Historic Data Through Mapping and Visualization Tools

Description: The Historical Society of Pennsylvania (HSP) provides its members with access to biographical data extracted from diverse collections of historical records. This data can be accessed by members through the Encounters platform: <http://hsp.org/collections/catalogs-research-tools/members-only-databases>. Using the Public Schools Admissions Records database, the Fellow will conduct research and normalize location data from the admissions records, so that data can be later mapped for visualization in Encounters.

Problems: The main challenge to be addressed in this project is the mapping of historic geo-spatial data to geographical coordinates in an automated fashion. Location data is inconsistently formatted, or may be missing. In addition, streets may have undergone name changes within the time period of the data, which means research is necessary to map historical addresses to modern-day coordinates. An additional challenge to be addressed is determining user-friendly tools that can be provided to users so that they may visualize connections between places and people.

Techniques: Techniques include data normalization; scripting; algorithms; and data visualization tools. Fellow will be working remotely with a MySQL database, using modules from an appropriate programming language, such as php, Perl, or Ruby. Mapping historical location data to modern coordinates may require a Google API key, which can be determined during the project. There are many resources available for these tasks; the key will be in determining what tools will be compatible with the Encounters platform.

Data: The information that has been transcribed and stored in database structure include: biographical data (e.g. life event dates, children and spouse names, occupational information), health data (e.g. cause of death, vaccination data); geo-spatial (locations of institutions and personal residences); and institutional data (e.g. data related to beneficial societies, public schools, government agencies, charitable organizations).

Outcome: Improved access to historic data for scholars and other user groups (e.g. family history researchers) via user-friendly visualization and mapping tools. The ability to visualize connections between people, places, and institutions.