Linking Multi-faceted Complex Data: From Vision to Reality

Jeanette Zerneke
Electronic Cultural Atlas Initiative (ECAI), School of Information, UC Berkeley

Linking and integrating data from scholars around the world and over time is a holy grail. The ability of integrate the productive work of scholars in a myriad of disciplines over millenniums will allow us to create a more seamless understanding of our world. What does it take to integrate complex multi-faceted information and build the ability to implement this vision?

Amazing tools are being developed to handle extremely large datasets of contemporary digital data and create a web of linked data. What we're more interested in is how to expand this network of data to include heterogeneous data and still enable complex visualization and analysis functions. The data we work with comes from many sources and has a large range of quality, accuracy, certainty and ambiguity.

How do you find and evaluate potential information to use, what is required to understand the pedigree of diverse data, and what does it take to manage and ingest the data so it can be used in an integrated environment? We have developed a matrix of the issues involved in sharing and re-use of data from diverse sources and a methodology for characterizing uncertainty and ambiguity. This process assists us in understanding how to maintain the pedigree information for the data so that future users of the data can fully understand the complexities of the information and how it can be used in an integrated visualization without implying information that isn't present.

In a case study of early California history I have implemented this methodology, informing the way that the information can be used in an integrated visualization and made available for future scholars and learners to integrate with their collections of data.