We present an open-source place finder. When developers want to put places on maps, or find potential duplicates between trees, they need to match the texts that genealogists write in “place” fields to standardized places. Finding the correct standardized place for text appearing in a place field is often challenging because people omit jurisdictional levels, enter ambiguous places, misspell place names, and sometimes append non-place information to places. The problem is made much harder for genealogical places because places change names and jurisdictional hierarchies over time.

Up to now there has not been an open-source place finder that people could incorporate into their software projects. This presentation reports on a newly-released open-source database of places that includes both historical and modern names and jurisdictional hierarchies, as well as geo-positioning information. It also reports on an open-source algorithm to match texts found in place fields to the standardized places in the database. The place matching algorithm uses machine-learning techniques to identify the most-likely place when matching place texts that are ambiguous.

The database of places comes from the community-supported database genealogical places at WeRelate.org. The algorithm was developed by matching nearly 8M place texts extracted from the 7000 GEDCOMs submitted to WeRelate over the past five years. The algorithm and database are both freely available on-line.

People can test the algorithm on-line by entering place texts into a server that we've launched. For more information see: https://github.com/DallanQ/Places

Authors:

Dallan Quass is the developer behind WeRelate.org, the world's largest public genealogy wiki. He has a PhD in computer science from Stanford University and BS and MS degrees from BYU. Prior to WeRelate, he was CTO for FamilySearch. Earlier, he co-founded WhizBang! Labs, FlipDog.com which was acquired by Monster.com, and Junglee which was acquired by Amazon.com.

Ryan Knight is a Software Architect with over 15 years of experience in all aspects of information technology, including enterprise architecture, technical sales, software development, project management, business analysis and database design. He currently runs his own consulting company where he has done a wide variety of projects such as: software architect for Deloitte, expert services for Adobe, a gift card creation tool for SendOutCards, and website creation for a variety of clients.