

## RESEARCH REPORT

Patrick Manning

### World-Historical Gazetteer

The CHIA project has completed initial steps in construction of a world-historical gazetteer. The objective is an ontological resource that will assist in documenting datasets with spatial information worldwide since 1500. The work is supported by an award from the Digital Humanities Start-up program of the U.S. National Endowment for the Humanities.

#### Workshop

Work began with a two-day workshop September 4-5, 2014, at which twenty scholars and developers—from numerous institutions, projects, and disciplines—met at the University of Pittsburgh to brainstorm and sketch out the objectives. Following reports by participants on current work— in software, infrastructure development, content, and organization – discussion turned to basic criteria for the gazetteer. It was agreed that a new and more general gazetteer, relying on open linked data, needs to be created for the period after 1500. This new resource is to emphasize simplicity in its basic form, and is to link smoothly to various and more specific spatial resources. It is to encompass places that

are specified with points and boundaries but also places for which we lack information on name, location, or feature type. The gazetteer is to allow specification of relationships among places, both hierarchical and parallel. It is to give special attention to coverage of Africa, Latin America, and Southeast Asia. The gazetteer project is to adopt, in large part, the strategy, tactics, and tools developed for pre-1500 gazetteers by Pelagios and Pleiades projects, in which Leif Isaksen, Rainer Simon, and Tom Elliott are key figures. It is also to work closely with the PastPlace project, led by Humphrey Southall, which is currently emphasizing creation of a worldwide list of administrative units.

### Comparison of geographic resources

As an initial follow-up to the workshop, Kathy Weimer (Rice University) and Tonia Sutherland (University of Alabama) conducted a preliminary survey of several geographic resources, to establish their relative strengths, especially in providing metadata to document place names. They gave special attention to spatial catalogues of the Library of Congress, Wikipedia, Geonames, and Getty. Their results indicated substantial overlap among the various resources—they cite each other in detail. All of them appear to be valuable but each has specific advantages over others. This investigation will continue in greater detail.

### Initial list of world-historical places

As an initial list of geographic terms, David Ruvolo (University of Pittsburgh) completed a scan of all the index terms in a leading historical atlas, *Atlas of World History*, edited by Jeremy Black. The index, scanned through an Optical Character Recognition program and copied to a large Excel file, yielded 14,000 entries, of which over 11,000 are place names. This is proposed as the core list for the world-historical gazetteer. This data has been organized in four ways: (1) a single column of over 11,000 place names; (2) a single column of 14,000 entries including place names and other entries (battles, peoples, cultures); (3) multiple columns, in which place names are accompanied by other descriptive data, including ID labels. These additional data also include information on the time period relevant to each place (since page numbers are linked to atlas chapters organized by time period).

In steps to come, the list of place names can undergo a process of disambiguation, especially to clarify duplicate names. The team of workshop participants will then decide on next steps to take in constructing the world-historical gazetteer.



Articles in this journal are licensed under a Creative Commons Attribution 4.0 United States License.



This journal is published by the University Library System of the University of Pittsburgh as part of its D-Scribe Digital Publishing Program and is cosponsored by the University of Pittsburgh Press.