

# World-Historical Gazetteer

Karl Grossner<sup>1</sup> and Ruth Mostern<sup>1,2</sup>

<sup>1</sup> University of Pittsburgh World History Center

<sup>2</sup> University of Pittsburgh Department of History

This paper reports on World-Historical Gazetteer (WHGazetteer), a three-year project funded by the US National Endowment for the Humanities, now at its midway point.

## Goals and Purpose

WHGazetteer is a scholarly infrastructure project intended to support historical research across many disciplines. It is principally a web-based software system for aggregating open access data about places and linking it with data about historical entities associated with those places. We have seeded the system with a global “spine” dataset of some 30,000 places, developed expressly for this purpose.

There are already existing repositories of historical place data, and many more are in development. A few are explicitly termed gazetteers<sup>1</sup>, some are historical GIS web resources<sup>2</sup>, and others comprise place data tables developed and published in the course of other research projects. Typically, each concerns a particular geographic area and a particular temporal scope<sup>3</sup>.

The WHGazetteer project, by contrast, is soliciting and aggregating attestations of historical places for *all* regions and periods along with annotations of records about historical entities with identifiers for those places. The project is conceived from a world-historical perspective. By this we mean several things. First, that it is intended to facilitate representations of connection and movement; second, that it scales up to global processes and the *longue durée*; and finally, that our “spine” data includes named ethnonyms, physical features, biomes, and ocean currents, helping ensure coverage for all parts of the globe and providing geographic context.

In many respects WHGazetteer follows and extends the pilot implementation of Peripleo<sup>4</sup>, the linked data gazetteer system developed by the Pelagios Commons<sup>5</sup>

---

<sup>1</sup> E.g. Pleiades (<http://pleiades.stoa.org>); Vision of Britain (<http://www.visionofbritain.org.uk>)

<sup>2</sup> E.g. China Historical GIS (<https://sites.fas.harvard.edu/~chgis/>), HGIS de las Indias (<https://hgis-indias.net/>)

<sup>3</sup> E.g. Map of Early Modern London (<https://mapoflondon.uvic.ca/>); Syriaca.org

<sup>4</sup> <http://peripleo.pelagios.org>

<sup>5</sup> <http://commons.pelagios.org>

project. We are working closely with the Pelagios group to leverage their accomplishments and to maintain the trajectory they established.

## Places and Traces

WHGazetteer will solicit contributions of two kinds of place-related data. One is attestations of place names found in historical texts and maps: Places. The other is annotations of records about any sort of historical entity—artifacts, creative works, persons, events, etc.—with place identifiers found in published gazetteers: Traces.

### *Places*

The WHGazetteer catalogues place references associated with a time period, which may be a date of publication for an historical source or temporal information about a name as in a modern historical atlas. In either case, a historical gazetteer records attestations from sources that assert that a name existed at a certain time. This approach to temporal scoping differentiates us from modern gazetteer data sources like GeoNames and Wikidata. A single place may (and ultimately should) have many linked attestations for different periods, where names, spellings, place type, relations, and geometry (location, size and shape) vary.

### *Traces*

Traces are assertions of spatial and temporal scope for historical items of almost any kind: the footprint of an individual's lifepath; the findspot of a coin hoard; the itinerary of a journey; the extent of a war; the places referred to in any sort of text, from a treaty to a novel; and so on.

## The WHGazetteer Ecosystem

### *Standards*

We are developing two standard contribution formats conforming to linked data requirements. The first, Linked Places format<sup>6</sup>, is essentially complete and is being tested with early contributions. The second, for Trace annotation data, based on the W3C Web Annotation Data Model, will be completed in early spring 2019.

### *Contribution Pipeline*

Contributors to WHGazetteer come in all sizes, with varying levels of technical capability. Some have elaborate web interfaces and maintain permanent URIs for thousands of place records. These projects will have little difficulty exporting abbreviated records transformed to Linked Places format. Others have dozens or at most hundreds of place references drawn from sources specific to their domain of interest and are unable to stand up per-place pages with permanent URIs. The

---

<sup>6</sup> <https://github.com/LinkedPasts/linked-places>

WHGazetteer contribution pipeline allows them to a) upload compatible CSV files, b) reconcile their records against Getty TGN, DBpedia, GeoNames and Wikidata as well as the WHGazetteer index itself, c) review and validate results of that matching process, thereby augmenting their records with *closeMatch* or *exactMatch* links, and d) contribute reviewed records to the index, published under permanent URIs provided by WHGazetteer.

### *Backend Stores and Middleware*

The WHGazetteer system backend is comprised of a PostgreSQL relational database and multiple Elasticsearch index stores. These and the APIs for internal use and public access are managed with Django, a Python-based web development framework.

### *Interfaces*

WHGazetteer will provide a public API supporting machine queries to all indexed data. A graphical web interface will support contribution activities for authenticated users, and provide search and mapping capabilities for both Places and Traces.

Linking Place and Trace data in a single backend allows us to present linked data portals for all indexed places, which will grow richer over time as our indexes expand. Using either the public API or graphical interface one might discover, for a given place: its names, shapes, and relations with other places over time; people whose lifepaths have intersected it; journeys for which it has been a waypoint; and texts and artworks for which it is a subject.

Our contribution pipeline interface will enable several kinds of pedagogical applications. For example, students and instructors will be able to create custom Traces associated with course material. Advanced students will be able to upload CSVs of gazetteers they have created, and use the WHGazetteer reconciliation tool to augment their records with information in the WHGazetteer index.

### *Status*

The Version 1 release of WHGazetteer is scheduled for January 2020. An early beta version will be accessible by July 2019.

We have established data partnerships with roughly a dozen future contributors of large datasets covering a range of regions and periods, and also have a queue of many smaller projects. Data pipeline functionality to receive these is nearing completion (e.g. Fig. 1).

### *Sustainability*

WHGazetteer has been designed to require minimal hand-holding for contributions and tools for efficient curation and maintenance. The World History Center is committed to maintaining the system for the foreseeable future.

World-Historical Gazetteer Maps API Contributing About WHG ▾ **Dashboard** Logout


**align\_tgn (alcedo)** task id: 9f54ac64-f867-4b77-8601-feb0e2607298

« first previous Record 15 of 7207 next last »

**Acapulco**

**Name variants:**  
Acapulco Save

**Place type(s):**  
city (ciudad)



flag **geometry** related closeMatch exactMatch  no match  
**tgnid:** 7430378  
**Names:** Acapulco;  
**Types:** inhabited places;  
**Parents:** Baja California, México, North and Central America, World

flag **geometry** related closeMatch exactMatch  no match  
**tgnid:** 7007127  
**Names:** Acapulco; Acapulco de Juárez;  
**Types:** transportation centers (inhabited places); inhabited places; resorts; trade centers; cities; ports (settlements); tourist centers;  
**Parents:** Guerrero, México, North and Central America, World

flag **geometry** related closeMatch exactMatch  no match  
**tgnid:** 7349775  
**Names:** Acapulco;

Figure 1 - Screen for reviewing and validating match results of reconciliation against Getty TGN (pilot)