Old Traces, New Links: Representation of Taiwan Baotu in OpenStreetMap

Jheng-Peng Huang [1,4], Hao-Syong Liu [2], Hsiung-Ming Liao [2], Tyng-Ruey Chuang [1,2,3]

[1] Institute of Information Science, Academia Sinica, Taiwan
[2] Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan
[3] Research Center for Information Technology Innovation, Academia Sinica, Taiwan
[4] Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, Taiwan

1. What Is Old Is New Again

There are various approaches to digital reproductions of old artifacts. Some emphasize a realistic presentation of the appearance of the artifact. One typical approach, for example, is to make a careful scan of a rare book in order to preserve faithfully the book's look-and-feel to human eyes in high-resolution images. Other approaches, however, aim to extract as much as possible information from the artifacts in order to be processed by machines. Such an approach will instead convert (manually, automatically, or by a mix of manual and semi-automatic procedures) the content of a book into pure text so that it can be indexed and text-mined by computers. The digital representation of an old book, in this case, is its computer text. Such a representation opens up opportunities for further reuse (e.g. audio books, hyperlinked pages, machine-translated scripts, etc.).

When working on old maps, researchers can take the latter approach: One can strike to retrace the contours and elements depicted in the maps, and convert them into interlinked digital objects. Old routes in ancient atlas become road networks in online maps supported by GIS (Geographic Information Systems). We can then make travel plans as if we were living in the past by using information extracted from the old maps. The routes, however, can be calculated with the help of computer algorithms. A well-known autography can be taken apart and its essentials re-created by repurposing modern day social network services, Facebook for example. Researchers and students can now indulge in an interactive chronology of the people, social circles, and events described in the book. The Web, with its generic facility for linking digital resources, allows us to construct new representations of old artifacts.

This short paper is a preliminary report of our experience in using OpenStreetMap to digitize and put online a small portion of Taiwan Baotu. OpenStreetMap is both a grass-rooted, collaborative effort in global map-making, and a reliable Web map service that is free for all to use [1, 2]. We modify and reuse the software behind the OpenStreetMap's Web map service, but apply it to geospatial data digitized from old maps. We hope what we do will come to exemplify a new approach to represent and

repurpose old maps; we aim to create online resources from ancient maps where old traces become new links. Further, the links can be collaboratively added and forged when such online resources are shared and reused.

Note that OpenStreetMap has also been adapted for other purposes. In many of these cases datasets from elsewhere are overlaid over OpenStreetMap, and it becomes convenient to browse them as themed annotations attached to a certain geographic area of interest. Often the online OpenStreetMap service is used as the source of the background map; this necessarily generates a modern map of the area [3]. Such an overlaying method is frequently applied to other web mapping services as well (Google Map/Earth, for example) so that online maps enriched with collections of cultural objects (in the forms of digital images, web links, online conversations, etc.) can be viewed and shared by many [4]. In this work, instead we hope to give new lives to old maps so that them can be the background maps, be interacted with, and be enriched with other datasets. We achieve this by modifying and reusing the OpenStreetMap server-side software package. For old maps that cannot be recoded by using modern coordinate systems, their representations in OpenStreetMap will be difficult, however, as OpenStreetMap depends on the use of precise geospatial coordinates. As such, our approach is accordingly limited in its applicability.

2. Why Taiwan Baotu?

The Taiwan Baotu is a set of 457 topographic maps covering a major part of the Taiwan island. Each map is 1:20000 scaled. It was produced in 1904 when Taiwan was under the Japanese rule. The collection was first published in 1906; another edition with redacted place names was released in 1920. Taiwan Baotu was derived from the output of an island-wide land survey, but had also incorporated other types of geographic information. The maps illustrate administration areas with their detailed boundaries. They contain place names, as well as land use, transportation, landmark, and other information about Taiwan in the early 1900s.

At Academia Sinica, Taiwan Baotu was previously scanned as high-quality images. These images are available as online services from which they can be used together with other map images. Like most Web-based map collections, they are delivered as map tiles to user browsers to be examined (zoom, pan, overlap, etc.). Intrinsic details in Taiwan Baotu, such as administrative boundaries and place names, are not available as machine-processable datasets. Parts of the Taiwan Baotu had also been digitized using GIS software package. The traced contours of the maps are stored as layers of vector data (in the Shapefile format, to be exact). As such, they can be visualized and analyzed by various GIS software (as long as the software support Shapefile), but such activities cannot be easily performed and coordinated over the Web.

As we have been participating in a multidisciplinary project about the Taijiang Inner Sea Area (Tainan, Taiwan), we aim to represent Taiwan Boutu as a web of interlinked resources upon which researchers can further enrich and collaborate with one another.

3. OpenStreetMap as Infrastructure

OpenStreetMap is a mass collaboration on mapping the world as it is now. We aim to learn from its success, and we want to use its technical infrastructure to represent the Taijiang Inner Sea Area as it was in the past. OpenStreetMap probably is best known for its practices and tools for collaborative map-making. In our work, we have not used much the collaborative part of the OpenStreetMap. Instead, the part we reply on most is its underlying sub-system for map rendering, tile serving, and overlapping. To represent Taiwan Baotu in OpenStreetMap, we first took as input an existing collection of Shapefile layers. We renewed, merged, and converted these layers into new datasets following the OpenStreetMap data format.

An OpenStreetMap system is built upon a collection of open source software packages: a database for storing map data, a rendering system for converting data into images (i.e., map tiles), and a tile server. The tile server responds dynamically to a user's request for the map of a certain area and of a particular scale. OpenStreetMap software is released under free/open source software (FOSS) licenses; everyone can freely use and modify the software to serve one's needs (as long as the software license is respected). This means the experiments we have with Taiwan Baotu with OpenStreetMap can be reproduced and validated by others; the experience can be easily shared in the research community (which is not true if proprietary software or service is used).

4. Taiwan Baotu in OpenStreetMap

Our current work focuses on three main tasks: Redrawing Taiwan Baotu, converting map data for use in OpenStreetMap, and providing an interactive layer on top of Taiwan Baotu in OpenStreetMap. These steps can be applied to any historical maps with definite geospatial coordinates.

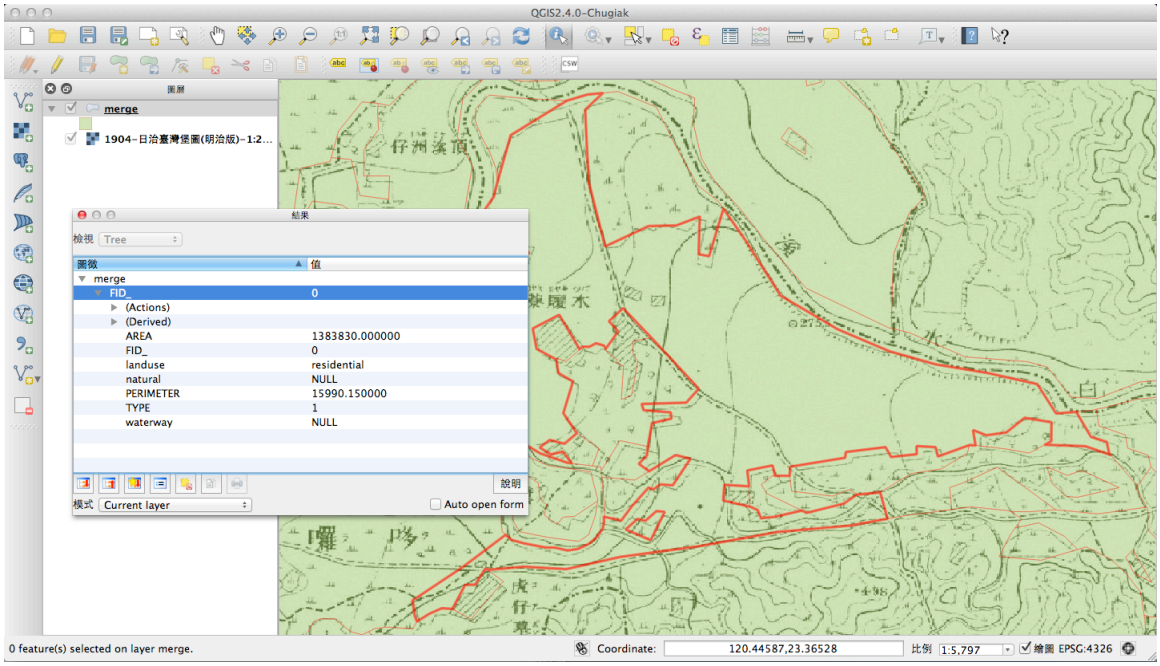

Figure 1. Redrawing maps in QGIS.

In the first step, QGIS is chosen as the tool to redraw the historical maps. Figure 1 shows a screenshot of the redrawing process by using QGIS. Note that in our case, most of the areas we are interested in had already been redrawn this way in QGIS. However, the data is of various quality. The redrawing is actually an iterative processing: Feedback from the OpenStreetMap end will trigger editing at the QGIS end, and the data modified with QGIS is again fed to OpenStreetMap for visualization and verification. The result from QGIS is saved in the ESRI Shapefile format which is a popular geospatial vector data format supported by most GIS software.

In the second step, we convert datasets from the Shapefile format to the OpenStreetMap format before they can be imported into OpenStreetMap. We developed a program to do the conversion automatically; we also provided a Web form for people to upload and convert Shapefile datasets so as to try to bridge the various gaps between the historians, the researchers, and the programmers. There are several existing conversion programs but they are not able to perform this job properly because our datasets contain Chinese characters (in various encodings). Our program supports Unicode and is able to import the result directly into the OpenStreetMap database after a map dataset is converted.

Note that once a Shapefile dataset is converted into the OpenStreetMap format, the dataset becomes readable and editable by simple text editors as the OpenStreetMap data format uses XML. XML is a preferable file format as it is a text format with many supporting tools. The OpenStreetMap rendering service will convert map datasets into map tiles. The tiles are sent by an OpenStreetMap tile server to display and interact h in user browsers.

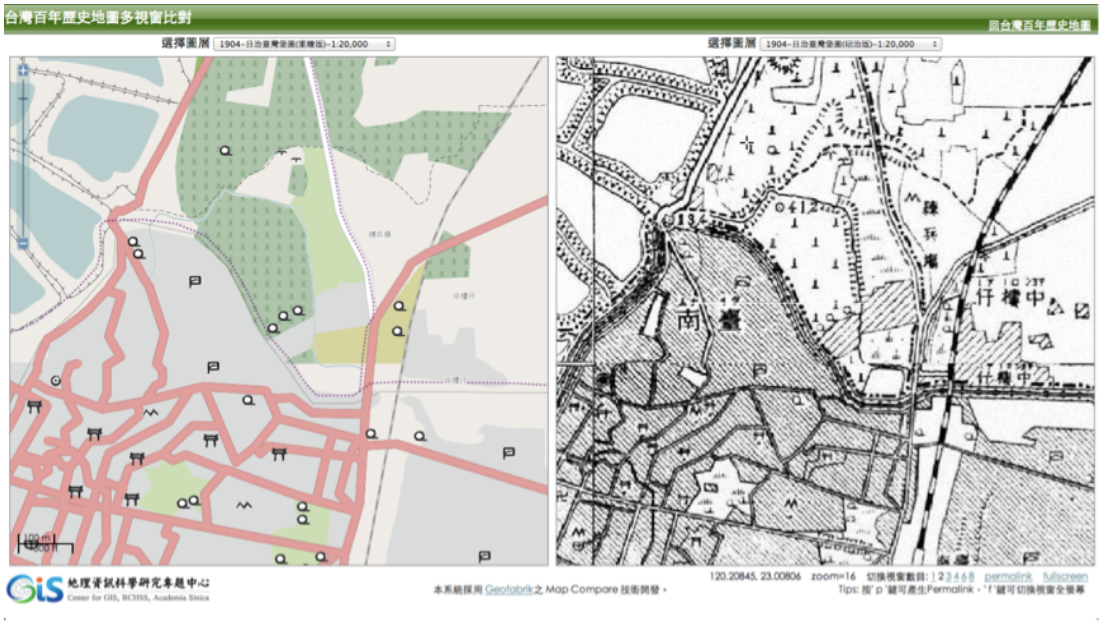

Figure 2. A comparison of a paper map and its digital representation in OpenStreetMap.

In the last step, we built an interactive map of Taiwan Baotu on which users can contribute data on top of the visualized map. Visualization work is usually supported by the use of a Javascript library. We use Leaflet to produce an interactive map to present Taiwan Baotu. Figure 2 shows two maps in the same screen. We use this platform to review the quality of the redrawn map: one can compare the difference between the paper map and the digital map. Leaflet can deal with GeoJSON data, which is a format for encoding a variety of geographic data structures, and shows the data on the map. Therefore, we used it to provide a platform for users to connect external data to the historical map on display.

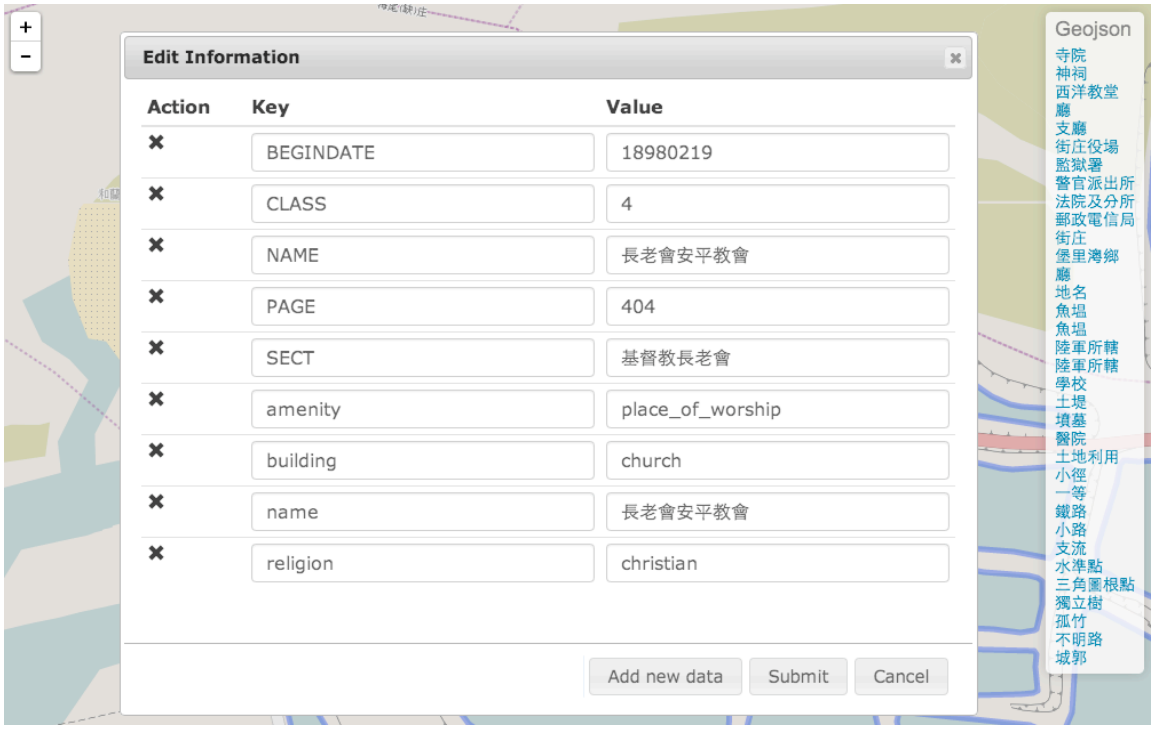

Figure 3. Attaching external information to a historical map.

The external data is usually in the form of a URL to an article, an image, or a video accessible over the Internet. We provide a form for users to fill in the URL of an external resource that is related to a location or an area on map, see Figure 3.

5. Openness in Research and Scholarship

Our method represents a workable and reproducible way for the digital representation of historical maps where intrinsic map details are kept and opened up for reuse. As we keep as much as information in the maps in digital and interactive forms, these historical maps become more helpful for the historians, researchers, teachers, students, and anyone who are interested in the maps and wish to contributed to the development of digital humanities. Recent research and observation about research data sharing and reuse (of which map but one category) shows that open scholarship is the norm. Formal and informal scholarly communication (sharing and reusing map data but one example) is converging, and open access to data is a paradigm shift [3]. Our methods are small but concrete steps to open up historical maps for better sharing and reuse.

References

1. Frederik Ramm, Jochen Topf, and Steve Chilton. *OpenStreetMap: Using and Enhancing the Free Map of the World*. UIT Cambridge, 2011.

2. Mordechai (Muki) Haklay and Patrick Weber, *OpenStreetMap: User-Generated Street Maps*. IEEE Pervasive Computing, vol.7, no. 4, pp. 12-18, October-December 2008.

3. Guillermo Amat, Javier Fernandez, Álvaro Arranz, and Angel Ramos. *Using Open Street Maps Data and Tools for Indoor Mapping in a Smart City Scenario*. In The 17th AGILE International Conference on Geographic Information Science, Castellón, June 2014.

4. Todd Samuel Presner, David Shepard, and Yoh Kawano. *HyperCities: Thick Mapping in the Digital Humanities*. Harvard University Press, 2014.

5. Christine L. Borgman. *Scholarship in the Networked World: Big Data, Little Data, No Data*. Oliver Smithies Lecture, Balliol College, Oxford, June 2013.