Marquette University
e-Publications@Marquette

GBU - Gazetteer of Beauplan's Ukraine   OpenOttoman

4-1-2018

GBU-Description

Mark Polczynski
Marquette University, mark.polczynski@marquette.edu

Michael Polczynski
Georgetown University, michael.polczynski@georgetown.edu

<u>**GBU: A Gazetteer of Beauplan's Ukraine**</u>

Michael Polczynski ~ michael.polczynski@georgetown.edu
Mark Polczynski ~ mark.polczynski@marquette.edu
May 10, 2018

Introduction

In 1630, Guillaume Le Vasseur, sieur de Beauplan, travelled to the Polish-Lithuanian Commonwealth to begin a seventeen-year military career in the Crown army. His account of this period was published in 1660 as *La description d'Ukraine*[1]. In addition to providing a coherent tableau of 17th century Ukrainian landscape, peoples, flora, and climate, Beauplan's work provides a major cartographic description of this region. The purpose of the *Gazetteer of Beauplan's Ukraine* (GBU) project is to provide a georeferenced database for places shown on Beauplan's maps. The primary output of this project is the *gazbu* database which captures over 3,000 populated places, rivers, river fords, river rapids, islands, forests, mountains, and valleys shown on Beauplan maps.

The purpose of this GBU project description is to: 1) enable users of the *gazbu* database to understand the structure and contents of the database; 2) help people new to the task of constructing a gazetteer to avoid common pitfalls and improve the efficiency and quality of their work. This description is structured as follow:

1. Overview of historical gazetteer creation;
2. Primary and modern data sources;
3. *gazbu* database fields;
4. Access and licensing of materials;
5. Beauplan map grid labeling;
6. Location and name confidence;
7. Data quality assurance;
8. Database linked data features;
9. Database creation process;
10. GBU project applications;
11. Related work;
12. Lessons learned.

1. Overview of Historical Gazetteer Creation

An overview of the general process for creating a gazetteer of historical places can set a context for discussion of the GBU project. As shown in Figure 1, creation of such a gazetteer involves six major elements. Of course, the first step is to secure sources of historical data, followed by selection of modern sources that will be used to geolocate historical places. These two elements are then use to extract geolocation and associated data about the historical places. In response to data extraction, a database structure is defined in a manner that accommodates the extracted data. The gazetteer database can then be populated, and the completed gazetteer can then be applied to answer questions and solve problems.

While this brief summary implies a sequential process as indicated by the arrows in Figure 1, in reality this is typically an iterative process with many false starts and much backtracking. In fact, the process is best started at the application element by "beginning with the end in mind". If the product this process does not accommodate anticipated applications, the entire exercise is essentially a waste of time and resources. Further, all of these elements are interdependent in that decisions and changes to each element can be expected to impact other elements. With these issues in mind, a discussion of the GBU project and *gazbu* database can now proceed.

2. Primary and Modern Data Sources

Entries in the *gazbu* database are derived from the Beauplan maps shown in Table 1. At the time of this writing, nine of the maps have been processed, with one work-in-process map. Codes shown in the table are the abbreviations for map names used for this project. The means of obtaining access to on-line digital versions of each map is described later in this document. For reference, Figure 2 shows an example of a Beauplan map, and Figure 3 shows a section of that map.

---

[1] https://www.worldcat.org/title/description-of-ukraine

Places on Beauplan maps are located through association with currently-existing places and other landmarks. Three sources of modern landmarks are used to make these associations. The authors' *Base Maps of Ukraine* (BMU) project[2] provides maps of rivers, terrain elevation[3], and terrain ruggedness[4] for the region covered by Beauplan's maps. The rivers base map in particular was useful in finding the general area where a Beauplan map place might be found. Exact locations for the majority of Beauplan places were obtained by associating the places with places in the GeoNames geographical database (http://www.geonames.org/), available for download free of charge under a creative commons attribution license. For Beauplan places that couldn't be associated with GeoNames places, the Google Physical map (https://www.google.com/maps/) was checked.

3. *gazbu* Database Fields

As described later, the *gazbu* database is available as a CSV file. The database fields are shown in Table 2. The first column in the table gives the field name as it appears in the CSV file. The second column designates the field format as either text or URL, the third column provides a brief description of the field, and the fourth provides an actual example taken from the *gazbu* database[5].

Per the *gazbu-entry* field in Table 2, every entry in the database has a unique ID given as a URL for a web page associated with the entry. Figure 4 shows the web page for the example of Table 2. Note that the field names of Figure 4 are links to web pages that provide the associated information given in Table 2. Figure 5 shows the web page for the *gazbu-entry* field of Figure 4.

The *gazbu-id* field of Table 2 gives the unique ID assigned to each place in the *gazbu* database. It is critical to note the difference between *gazbu-entry* and *gazbu-id*. A particular place may appear on multiple maps with multiple spellings on each map. Such a place will have multiple entries in the database with each entry having a different *gazbu-entry* value, but all such entries will have the same *gazbu-id* since all entries are associated with the same place.

An entry's *gazbu-map* value contains the URL link to an on-line digital version of the particular map on which the entry can be found. Figure 6 shows the web page for the map with map code *dgu* (see Table 1 for map codes). Clicking on the map source URL of Figure 6 opens an on-line digital version of the map.

The *gazbu-grid* field helps users find places on *gazbu* maps. In Table 2, database entry 12 can be found in grid square *R12* on the *ds2* map. The means for generating grid squares on maps is explained later in this document.

Places on Beauplan's maps are divided into eight *gazbu-class* values: *frd*=ford, *fst*=forest, *isl*=island, *mnt*=mountain, *pop*=populated place, *rap*=rapids, *riv*=river, *val*=valley. Per Table 2, database class values are URLs for web pages that provide a description of each class. Figure 7 shows the web page for class=*riv*.

The *gazbu-name* field contains an entry's name as spelled on a map. Confidence in the spelling of the name is captured in the *gazbu-name-confidence* field, with field values being URLs that link to confidence level descriptions. Rules for assigning confidence levels are provided later in this document.

Comparing the *gazbu-name* and *latin-name* example values in Table 2 illustrates how special characters are changed in Latinized versions of names. The database contains Latinized spellings to facilitate database searches.

As mentioned in the previous section, *gazbu* database entry locations are taken primarily from associated GeoNames database places, with locations being entered into the *gazbu-latitude* and *gazbu-longitude* fields. The GeoNames database URL for the place appears in the *geonames-id* field.

The *gazbu-location-confidence* field reflects the strength of the association between the *gazbu* place and the GeoNames place, with the rules for assigning confidence levels being provided later in this document. Again, field values are URLs linking to descriptions of each confidence level.

---

[2] https://epublications.marquette.edu/ottoman_ukraine/
[3] Shuttle Radar Topography Mission data (http://lta.cr.usgs.gov/SRTM1Arc) was used to produce the BMU project terrain elevation map.
[4] Riley, S. J., S. D. DeGloria and R. Elliot, A terrain ruggedness index that quantifies topographic heterogeneity, Intermountain Journal of Sciences, vol. 5, No. 1-4,1999 (https://download.osgeo.org/qgis/doc/reference-docs/Terrain_Ruggedness_Index.pdf). The terrain ruggedness index reflects the difference in elevation values from a center cell to the eight cells immediately surrounding it.
[5] As a point of clarification, the thirteen field names of Table 2 are the column headers in the *gazbu* database, with each entry in the database being a separate row in the database. Thus the example column in Table 2 provides an example of one row in the database.

The *contributor-id* field contains a URL for the contributor of an entry. Ideally, this field contains an e-mail address or some other means for contacting the contributor.

It is worth noting that following the *geonames-id* URL link for a particular place opens a web map showing the associated GeoNames place. Figure 8 shows the web map page for the example of Table 2, where GeoNames give the modern name of this river as *Sukhyy Tashlyk*. Note also that the locations of all of the types of places included in the database are given as a single latitude/longitude pair located at the approximate center of the place, but river locations are the latitude and longitude of the mouth of the river. Because of this, it can happen that the location of a river shown on a particular Beauplan map may not lie on the map if the river's mouth lies off the map. A similar situation can occur if the center of a forest shown on a particular map lies off of the map.

4. Access and Licensing

The *gazbu* database is available as a .csv file in Unicode UTF-8 text format. The database can be viewed as a web map at https://qgiscloud.com/polczynski/GBU_Web_Map/ . All GBU materials are available through the project repository[6]. All GBU-related materials are provided through a Creative Commons Attribution 4.0 International (CC BY 4.0) license[7]. Per the license, you may copy and redistribute the material in any medium or format, and may remix, transform, and build upon the material for any purpose. If you do so, you must give appropriate credit by referencing the *gazbu* project repository[6] and indicate if changes were made to the materials. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use. .

5. Beauplan Map Grid Labeling

The primary task in creating an entry in a gazetteer such as the *gazbu* database involves going from a place on a historical map to a place on a modern map, but actually using the information in the database may involve going from an entry in the database back to the associated place on a historical map. The *gazbu-grid* field in Table 2 provides the ability to locate a database entry on a Beauplan map. Beauplan maps typically include latitude and longitude markers along the edges of the map. A map grid is created for a particular map by drawing grid lines between these markers. Using the markers actually shown on a Beauplan map allows anyone with access to the Beauplan map to re-create grid squares that correspond to the *gazbu-grid* values in the *gazbu* database.

Figure 9 shows the upper left corner of Beauplan's *dgu* map with grid lines drawn in this manner. Grid squares are labeled with letters and numbers as shown in the figure. Here, the place named *Temruk* is located in grid square *W15*. Table 3 shows the latitude and longitude grid square labeling system used for all maps included in the GBU project. Per the table, grid square *W15* lies between longitudes 59º and 60º, and between latitudes 45º and 46º. Note, however, that latitude and longitude markers on Beauplan maps are not consistent or geographically accurate. Figure 10 shows *Temruk* on Beauplan's *tgu* map, which falls in grid *Q14* on this map. Note also that while projecting actual lines of latitude onto a flat map produces curved lines, GBU grid lines are drawn as straight lines between the latitude and longitude markers shown on Beauplan's maps.

6. Location and Name Confidence

Places on Beauplan maps are geolocated by association with modern places found in the GeoNames database. Location confidence (*gazbu-location-confidence* in Table 2) reflects the strength of the association. In general, location confidence levels of 1 through 5 are assigned as follows[8].
1. Strong association between historical place name and modern place name.
2. Strong association between historical place name and modern place name but several modern places in the area with some form of name qualifier (example provided later in this document).
3. Weak association between historical place name and modern place name.
4. No association between historical place name and modern place name but strong association to modern place with different name based on surrounding landmarks.
5. No association between historical place name and modern place name and weak association with other surrounding landmarks.

---

[6] https://epublications.marquette.edu/ottoman_beauplan_gazetteer/
[7] https://creativecommons.org/licenses/by/4.0/
[8] Note that as both location and name confidence appear as URLs in the *gazbu* database, the confidence level of an entry appears as the numeric value in the URL. So, an entry with *http://gazbu.org/location-confidence/1.htm* has a location confidence level of 1.

For all confidence levels except level 5 the GeoNames ID and location of the associated modern place are assigned to the historical place. For level 5 the GeoNames ID value of *http://www.gazbu.org/geonames-id-none/* is assigned to the place and location is approximated based on surrounding landmarks.

Regarding location confidence level 2, an example would be two places with the same name except one name has a qualifier like *stary* (old) and the other has a qualifier like *nowe* (new). If surrounding landmarks do not indicate a preference for one place, the place with the more likely qualifier is selected (here *stary* vs. *nowe*).

In constructing the database, there were places on Beauplan maps that could not be associated with places in the GeoNames database but that could be associated with places shown on the Google Physical map. Such places were added to the GeoNames database[9] at the location shown on the Google Physical map, thereby providing a GeoNames ID for the place. It should be noted that GeoNames and the Google Physical map sources name primarily populated places and rivers and often do not name fords, forests, islands, mountains, rapids, and valleys, so most of the *gazbu* entries for these types of places have no GeoNames ID and a location confidence = 5 even if surrounding landmarks provide a good approximation of the location of these places.

An exception to this scheme for assigning location confidence applies to rivers shown on Beauplan maps that can be associated with rivers that are shown on the Google Physical map but not named by either the Google Physical map or GeoNames. Here, location confidence levels 1 and 3 are as follows:
1. Strong association between historical river and un-named Google Physical map river. GeoNames ID = *http://www.gazbu.org/geonames-id-none/* and location = Google Physical map river mouth location.
3. Weak association between historical river and un-named Google Physical map river. GeoNames ID = *http://www.gazbu.org/geonames-id-none/* and location = Google Physical map river mouth location.

Table 4 shows statistics on *gazbu* database location confidence levels, with the difference between total and unique counts reflecting the occurrence of some places on multiple maps. The statistics indicate that about 4 out of 5 places shown on Beauplan maps can be associated with currently-existing places.

Regarding name confidence, Beauplan name values in the *gazbu* database (*gazbu-name* in Table 2) are spelled as shown on the Beauplan maps, but the names are not always legible especially with regard to special characters such as ÿ and ł. Name confidence levels of 1 through 5 reflect the legibility of the name.
1. All letters in name clearly legible.
2. Not certain if/which special characters are used in the name.
3. One or two letters in the name not clearly legible.
4. Name partially legible.
5. Name not legible.

Note that different Beauplan maps may spell names of the same place differently, and since each name spelling produces a separate entry in the database, different entries for a particular place may have different name confidence levels.

7. Data Quality Assurance

Being a manual process, entering values into the *gazbu* database is error-prone. In order to improve the quality of the data within the database, entries are tested per the rules of Table 5 and entry corrections are made accordingly.

8. Database Linked Data Features

A major challenge faced when constructing a gazetteer is that a particular place may be found on multiple sources at different locations with multiple names and name spellings. The problem is exacerbated for gazetteers of historical places, since all of these place attributes typically change over time. Attempting to capture such many-to-many relationships can result in a complex database structure that is difficult to maintain and expand. One solution is the use of a linked data structure[10,11].

The GBU project attempts to incorporate certain aspects of linked data principles into the *gazbu* database. A highly-recommended linked data practice is to use URLs for data values wherever possible. Per Table 2, the *gazbu-entry*, *gazbu-map*, *gazbu-class*, *gazbu-name-confidence*, *gazbu-location-confidence*, and *geonames-id* fields contain

---

[9] For how to add places to the GeoNames database, see http://www.geonames.org/manual.html#create .
[10] http://linkeddata.org/
[11] http://www.w3.org/standards/semanticweb/data/

links which provide additional information relevant to an entry. For example, the *gazbu-map* value links to an on-line digital map source per Figure 6.

The way in which the *gazbu* database accommodates many-to-many relationships is illustrated in Table 6. Here, place *gazbu-514* is found on two maps in different map grids, with two different names on each map. Place *gazbu-516* is found on two maps within different map grids and with different spellings of the same name. In practice, a search of the database for *Podhaicze* would return all fields from *gazbu-entry* number 3006, which has a *gazbu-id* value of *gazbu-516*. A follow-on search for *gazbu-516* would then return data for *gazbu* entries 3006 and 1176, thereby revealing the two spellings and two map sources and map grids for the place.

9. Database Creation Process

Given this overview of the GBU project, it is possible to provide a brief summary of the process used to create the *gazbu* database. The primary tool used to create the database was the QGIS geographic information system (https://qgis.org/). Figure 11 shows a QGIS project screenshot illustrating how this tool was used.

The *Layers Panel* of the project has a *gazbu-SHP* shapefile layer that contains the places shown on Beauplan maps that have been geolocated, a *GeoNames* layer showing places in the GeoNames database to which Beauplan places are associated, a *BMU-Rivers-SHP* layer and a *BMU-Elevation-TIF* layer from the *BMU* project to help loate Beauplan places, and a *Google Physical* layer to locate places not found in the GeoNames database (the elevation and physical layers are not how for purposes of illustration).

The figure provides an example where the Beauplan place named *Sosenka* has been associated with the GeoNames place named *Sosonka* and assigned a location confidence level of 1. Also, the place spelled *Priluka* on one Beauplan map and *Prziluka* on another Beauplan map has been associated with the GeoNames place named *Pryluky*. Here, GeoNames give *Pryluky* an alternate name of *Priluka Staraya*, and also shows a different place named *Nova Pryluka*. Per the preceding definitions of location confidence, the GeoNames place with alternate name *Priluka Staraya* was chosen as the location of the Beauplan place and the Beauplan place was assigned a confidence level of 2.

Figure 12 shows a section of the completed geolocation process with the *gazbu-SHP* layer, the *BMU-Rivers-SHP* layer, and the *Google Physical* layer displayed. Note how location confidence is signified on the map. After geolocation is complete, the *gazbu-SHP* layer is saved as a .csv file, thereby creating the final *gazbu* database.

Per Table 1, there remain a number of Beauplan maps that will be added to the GBU gazetteer. Some mention of how this can be done in an efficient and reliable manner is in order. One approach is to use a spreadsheet program like Excel or LibreCalc to strip out all of the duplicate places in the *gazbu* database (places with the same *gazbu-id*) and then execute the process outlined above using this stripped-down *gazbu* database in place of the GeoNames database (call the sripped-down database *gazbuX*). When a place on the new map can be associated with a place in the *gazbuX* database, the *gazbuX* database entry for the place can have the *gazbu-entry*, *gazbu-map*, *gazbu-grid*, *gazbu-name, gazbu-name-confidence*, *latin-name*, and *contributor-id* fields updated as needed per the map being added. When all such associations have been made, the spreadsheet program can be used to strip out all of the *gazbuX* entries that do not belong to the new map, and the remaining places on the new map can then be located per the preceding process using the GeoNames database and Google Earth. Finally, the *gazbuX* database (which contains places only on the new map) can be appended to the *gazbu* database, and the quality tests of Table 5 can be applied.

While on the topic of processing additional maps, it can happen that different primary sources show a particular place in different locations such that that the place is associated with different modern places. If it is clear that the two places on the primary sources are indeed the same place, the place should be entered into the database with the appropriate database fields for both instances, except that both entries should have the same *gazbu-id* value. Since the *gazbu* database includes contributor information, ideally such apparent conflicts would be addressed by the contributors of the entries.

10. GBU Project Applications

In general, databases have no intrinsic benefit – benefit is derived from application of a database to answer questions and solve problems. To-date the *gazbu* database has supported two such applications: the BSZLAK project[12] which traces trails across the Pontic Steppe in the early modern period, and the SKBD[13] project, which provides a geographical context for the 1539-1542 border dispute between the Polish/Lithuanian Commonwealth and the Ottoman Empire.

---

[12] https://epublications.marquette.edu/ottoman_bszlak/
[13] https://epublications.marquette.edu/ottoman_skbd/

Szlak, a Polish word meaning "trail", was the name given to paths used to traverse the Pontic Steppe in the early modern period. Regular roads were virtually non-existent, and overland travel tended to follow general paths that conducted travelers between destinations as expeditiously as possible. The BSZLAK project attempts to trace the paths of three of these trails: the *Czarny szlak*, the *Kuszmanski szlak*, and the *Moraski szlak*, which were of particular importance to the Black Sea slave trade that saw the abduction or death of some two million individuals. These trails are shown on several of the Beauplan maps used to generate that *gazbu* database, thereby providing landmarks that allow trail paths to be traced. Figure 13 shows a section of a Beauplan map with the intersection of two portions of the *Czarny szlak* (with north at bottom), and Figure 14 shows the path traced on a modern map using populated places from the *gazbu* database.

The SKBD project provides insight into the Sawran-Kodyma border dispute. In 1539 a proposal for a joint border demarcation commission to establish the border between the Ottoman Empire and the Polish-Lithuanian Commonwealth was sent by Ottoman sultan Süleymân to Polish king Zygmunt I. By 1542, the issue had boiled down to a dispute over whether the border should extend between the Dniester river and the Pivdennyy Buh river along the Sawran river (Ottoman proposal) or along the Kodyma river (Polish proposal). While detailed period text accounts of the negotiations exist, the purpose of this work is to expand understanding of the situation by addressing questions such as: Where are the Dniester, Pivdennyy Buh, Sawran, and Kodyma rivers? How much territory lies in the disputed area? Were there any settlements, travel paths, river fords, forests, or other resources of importance in the disputed area? Ultimately, was this a struggle for control of a strategic resource-rich territory, or fundamentally a war of wills over a worthless plot of land? Figure 15 shows a section of a Beauplan map covering the disputed territory, and Figure 16 shows the equivalent region using places from the *gazbu* database and the Kuszmanski szlak from the BSZLAK project superimposed on rivers and terrain base maps from the BMU project.

11. Related Work

The GBU is the initial project of the Ottoman Historical Gazetteer (OHG), which is being developed within the OpenOttoman initiative. OpenOttoman[14] is a developing transnational digital collaborative for scholarship and public history focused on the Ottoman world. The OpenOttoman mission is to create new co-operations that cross geographical, linguistic, political, ideological and disciplinary boundaries that have, in the past, placed restrictive barriers on research in Ottoman studies. OpenOttoman seeks to enhance and sustain research, the learning experiences of students, and the availability of scholarship to a broader public, based on a commitment to be open, inclusive and accessible. Its goal is to leverage the possibilities offered by digital platforms and formats in order to enable scholars to connect with each other with published work and with research resources relevant to Ottoman studies.

The OHG was a collective decision of participants at the Digital Ottoman Platform workshops that took place at the Institute for Advanced Study (Princeton) in summers 2015 and 2016.[15, 16] There was general agreement that such a geo-referenced catalogue of places would constitute the greatest contribution to the largest number of scholars, both within the vast spatial reach of Ottoman studies and beyond its chronology and topography to areas connecting with its realms. In response to the need for an Ottoman world gazetteer, the GBU team created the Historical Gazetteer of Crimea[17] (HGC) based on Henryk Jankowski's historical-etymological dictionary of pre-Russian habitation names of the Crimea[18] as a pilot project and use case for further work. The GBU project represents a next-step effort with a level of gazetteer magnitude and sophistication beyond the HGC.

12. Lessons Learned

Generation of the *gazbu* database has been the result of several years of work, with frequent backtracking and multiple false starts. The following is an accumulation of lessons learned that may be of assistance to others attempting such an effort.

When creating a gazetteer database, it can be tempting to dive right in and start geolocating populated places, but for the GBU project it was highly beneficial to have a base map of rivers available when locating populated

---

[14] https://openottoman.org/

[15] https://www.ias.edu/ideas/2015/singer-digital-ottoman/

[16] https://www.hs.ias.edu/sites/hs.ias.edu/files/Schmidtke/DOP_II_Program%20and%20Abstracts.pdf

[17] https://epublications.marquette.edu/ottoman_crimea/

[18] https://www.worldcat.org/title/historical-etymological-dictionary-of-pre-russian-habitation-names-of-the-crimea/oclc/803613133

places, since most historical populated places were sited on rivers. Fords, rapids, islands, forests, mountains, and valleys can then use rivers and populated places as location landmarks.

Rubber sheeting is the process of locating places on a map by virtually stretching the map over another map with known place locations so that a few landmarks on the first map are aligned with associated places on the second map, thereby allowing the remaining locations on the first map to be derived from associated locations on the second map. For the GBU project, attempts were made to rubber sheet Beauplan maps over the Google Earth Physical map, but Beauplan map distortions and inaccuracies made this approach to locating Beauplan places unfeasible. Instead, places were located one-by-one "the hard way" starting with a rivers base map.

The overall gazetteer database generation process is iterative. After the first pass through place location, it is expected that the location of some places will have a confidence level less than 1. Here, iteration through the process using landmarks created in the previous pass may change the location of the place and/or increase the place's location confidence. At least two passes through the process are necessary to produce high-quality results.

Per the recommendation to start with locating rivers, the rivers shape file included in the BMU project was an essential tool in generating the *gazbu* database. Unfortunately, the BMU project was completed after the *gazbu* project was started and before the importance of rivers to the overall geolocation process was realized. In light of this lesson learned and in conformance with the iterative nature of database generation, rivers were added to the BMU project rivers shape file over the course of the *gazbu* project as examination of Beauplan's maps revealed rivers useful for locating places but not previously included in the BMU project.

When creating a gazetteer using historical sources, place name spellings often differ from modern spellings. This significantly limits the effectiveness of name text matching when searching for associations between historical and modern places. However, the actual pronunciation of historical and modern names can be quite similar, so when searching for a match between historical and modern places it can be beneficial to actually speak out loud the place names. For example, Beauplan's *Ciećeref* river is the modern *Teteriv* river, which sound similar when spoken out-loud. Of course, this is facilitated by some knowledge of how words are pronounced in different languages.

As previously noted, names on historical and modern sources may include a qualifier. For example, rivers on Beauplan maps may include the word *sucha*, which can be translated as *dry*, *mokra*, which is *wet*, and *dolina*, which is *valley*. Knowing the meaning of these qualifiers as well as how words are pronounced can assist in associating historical places with modern places. A complication factor for Beauplan maps in particular is that names are spelled in a mixture of Latin, French, and whatever was the language of the residents of the place at the time that Beauplan drew his maps. This can become a challenge when attempting a dictionary look-up of the meaning of a name or name qualifier.

While understanding languages and pronunciations can be of great assistance when generating a gazetteer, a person with limited understanding of historical and/or modern languages can identify patterns that assist in associating places. A common historical-to-modern change for Beauplan maps is *ow* to *'kiv*, and the letters *i* and *y* are often repeated and/or exchanged with abandon in both historical and modern place names. But as noted previously, the pronunciation of names with these characteristics is often quite similar.

When executing a multi-year project requiring several iterations through the database generation process, changes and constant additions are to be expected. This situation is exacerbated when several people are contributing to a database, or where different databases are integrated into a single database. One way to reduce database synchronization problems is to exclude data that is reliably stored in some other database. For example, the *gazbu* database does not include a field for modern primary and alternate names. This is because the *gazbu* database includes a link to the GeoNames database (*geonames-id*) which contains modern primary and alternative names for places, thereby eliminating the need and complication of including and updating modern primary and multiple alternative names in the *gazbu* database.

Finally, while this may seem a bit trivial, when executing a project with the magnitude and complexity of the GBU project it is essential to create and adhere to a descriptive file naming scheme and logical folder structure for the multitude of folders, files, and file types that get generated as part of the gazetteer construction process. Given the amount of effort required to construct a reliable and robust gazetteer, hunting for, loosing, and/or accidentally deleting files is frustrating and discouraging.

Conclusion

It has taken over three years to generate a gazetteer of 4,000+ entries for places appearing on nine Beauplan maps, with an additional two years of work anticipated for the remaining maps of Table 1. Our work is extremely modest compared to efforts such as Pleiades[19], Pelagios[20], the World-Historical Gazetteer[21] and the Cultures of Knowledge[22] project, but we hope that our efforts can benefit other scholars and welcome the opportunity to share the fruits of our labors.

While essential to execution of ongoing GBU team work, the *gazbu* database currently has limited opportunity to include and benefit scholars beyond the team. To this end, future GBU team efforts will explore means of providing better and wider access to project materials. For updates on our efforts, check our ResearchGate project log[23] and the GBU project repository[6]. For current and future related projects, see the GBU team's OpenOttoman repository[24].

Table 1. Beauplan maps used to construct the *gazbu* database.

Status	Code	Date	Beauplan Map Name
-	bpb	1665	Basse Podolie ou Palatinat de Braclaw
-	bpk	1665	Basse Volhynie ou palatinat de Kiow
-	cdu	1686	Carte d'Ukranie Contenant plusieurs Provinces comprises entre les Confins de Moscovie
Done	dgu	1648	Delineatio generalis Camporum Desertorum vulgo Ukraina
Done	ds1	1650	Delineatio specialis et accurata Ukrainae - Dzikie Pole - Panel 1
Done	ds2	1650	Delineatio specialis et accurata Ukrainae - Czerkasy Kaniow - Panel 2
Done	ds3	1650	Delineatio specialis et accurata Ukrainae - Kamieniec Podolski - Panel 3
Done	ds4	1650	Delineatio specialis et accurata Ukrainae - Halicz - Panel 4
Done	ds5	1650	Delineatio specialis et accurata Ukrainae - Putywl Baturyn Łochwica - Panel 5
Done	ds6	1650	Delineatio specialis et accurata Ukrainae - Kijow - Panel 6
Done	ds7	1650	Delineatio specialis et accurata Ukrainae - Zytomierz Berdyczow - Panel 7
Done	ds8	1650	Delineatio specialis et accurata Ukrainae -Lwow Luck Tarnopol - Panel 8
WIP	hpk	1650	Haute Podolie ou Palatinat de Kamieniec
-	hvl	1665	Haute Volhynie ou Palatinat de Lusuc
-	npl	1630	Nova totius Regni Poloniae Magniq. ducatus Lithuanae
-	pol	1702	La Pologne
-	ppl	1677	Le Royaume de Pologne Comprenant les Etats de Pologne et de Lithuanae
-	rnp	1665	Russie Noire divisee en ses Palatinats
-	rpl	1739	Regni Poloniae magnique ducat Lithuaniae nova
-	tgu	1648	Typus Generalis Ukrainae sive Palatinatuum Podoliae Kioviensis et Braczlaviensis
-	ubp	*	Ukrainae pars quae Barclavie palatinatus vulgo dicitur
-	ukp	*	Ukrainae pars quae Kiovia palatinatus vulgo dicitur

[19] https://pleiades.stoa.org/

[20] http://commons.pelagios.org/

[21] http://whgazetteer.org/

[22] http://www.culturesofknowledge.org/

[23] https://www.researchgate.net/project/Gazetteer-of-Beauplans-Ukraine

[24] https://epublications.marquette.edu/ottoman/

Table 2: *gazbu* database fields.

gazbu Database Field	Type	Description	Example
gazbu-entry	url	Unique ID assigned to each entry in the gazbu database	http://gazbu.org/entry/12.htm
gazbu-id	text	Unique ID assigned to each place in the gazbu database	gazbu-10
gazbu-map	url	Map on which this gazbu database entry can be found	http://gazbu.org/map/ds2.htm
gazbu-grid	text	Map grid on gazbu map where this entry can be found	R12
gazbu-class	url	Type of place as shown on gazbu map	http://gazbu.org/class/val.htm
gazbu-name	text	Name as spelled on gazbu map	Krywy Taszłeys
gazbu-name-confidence	url	Confidence of gazbu name spelling	http://gazbu.org/name-confidence/2.htm
latin-name	text	gazbu name with Latin characters	Krywy Taszleys
gazbu-latitude	text	gazbu place latitude	48.34947
gazbu-longitude	text	gazbu place longitude	30.89544
gazbu-location-confidence	url	Confidence of gazbu latitude and longitude	http://gazbu.org/location-confidence/1.htm
geonames-id	url	GeoNames ID of associated gazbu place	http://www.geonames.org/830038/
contributor-id	url	ID of person contributing gazbu entry	mhp.techforge@gmail.com

Table 3: Latitude and longitude grid square labeling scheme.

Longitude	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
From	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62
To	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63

Latitude	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
From	59	58	57	56	55	54	53	52	51	50	49	48	47	46	45
To	60	59	58	57	56	55	54	53	52	51	50	49	48	47	46

Table 4: *gazbu* database location confidence statistics.

Class	Total	Unique	Location Confidence (Unique)				
			1	2	3	4	5
Ford	45	35	0	0	0	0	35
Forest	59	44	6	0	2	0	36
Island	32	25	1	0	4	0	20
Mountain	5	5	1	0	1	0	3
Populated Place	3455	2542	1929	145	182	160	126
Rapids	26	15	0	0	0	0	15
Rivers	513	345	240	10	69	26	0
Valleys	23	12	2	1	0	0	9
TOTAL	4158	3023	2179	156	258	186	244
		CUMULATIVE %	72%	77%	86%	92%	100%

Table 5: Gazetteer data quality tests.

Test	Rule
duplicate-entry	If two entries have the same values for all fields except *gazbu-entry* Then these are duplicate entries
gazbu-entry	If two entries have the same *gazbu-entry* Then these are duplicate *gazbu-entry* values
gazbu-id	If two entries have the same *gazbu-id* Then they should have the same *geonames-id* and *gazbu-class* and *gazbu-latitude* and *gazbu-longitude*
geonames-id	If two entries have the same *geonames-id* except *geonames-id-none* Then they should have the same *gazbu-id* and and *gazbu-class* and *gazbu-latitude* and *gazbu-longitude*
gazbu-location	If two entries have the same *gazbu-latitude* and *gazbu-longitude* and *gazbu-class* Then they should have the same *gazbu-id* and the same *geonames-id*
gazbu-name	If two entries have the same *gazbu-name* Then they should have different *gazbu-map* or *gazbu-class* or *gazbu-latitude* or *gazbu-longitude*
gazbu-grid	If two entries have the same *gazbu-ID* and *gazbu-map* Then they should have the same *gazbu-grid* and same *gazbu-location-confidence* .
gazbu-GeoNames	If *geonames-id* is same as GeoNames database ID Then *gazbu-latitude* and *gazbu-longitude* should match the GeoNames database latitude and longitude

Table 6: Segment of *gazbu* database with many-to-many relationships.

gazbu-entry	gazbu-id	gazbu-map	gazbu-grid	gazbu-name
http://gazbu.org/entry/1174.htm	gazbu-514	http://gazbu.org/map/dgu.htm	M11	Parnasse
http://gazbu.org/entry/1175.htm	gazbu-514	http://gazbu.org/map/dgu.htm	M11	Podorze
http://gazbu.org/entry/3910.htm	gazbu-514	http://gazbu.org/map/ds8.htm	M10	Parnasse
http://gazbu.org/entry/3911.htm	gazbu-514	http://gazbu.org/map/ds8.htm	M10	Podorze
http://gazbu.org/entry/1176.htm	gazbu-516	http://gazbu.org/map/dgu.htm	M11	Podhÿce
http://gazbu.org/entry/3006.htm	gazbu-516	http://gazbu.org/map/ds4.htm	L11	Podhaicze

Figure 1: Creating a gazetteer of historical places.

Figure 2: Beauplan's *Typus Generalis Ukraine* map.

Figure 3: Section of *Beauplan's Typus Generalis* Ukraine map.
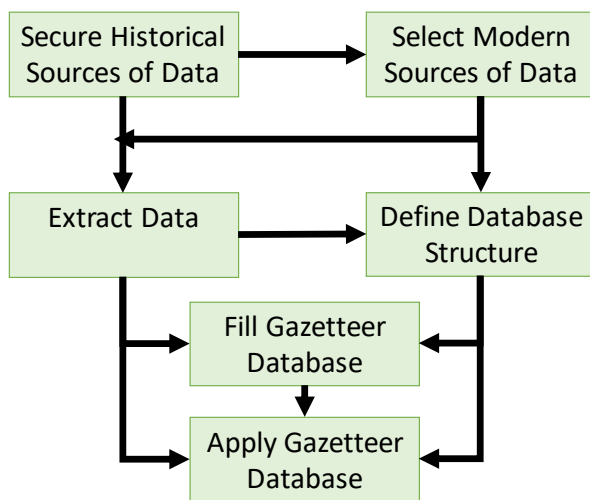
Figure 4: *gazbu* database entry web page.

Gazetteer of Beauplan's Ukraine
gazbu Database Entry Values
gazbu-entry: http://gazbu.org/entry/12.htm
gazbu-id: gazbu-3028
gazbu-map: http://gazbu.org/map/ds2.htm
gazbu-grid: R12
gazbu-class: http://www.gazbu.org/class/val.htm
gazbu-name: Krywy Taszłeys
gazbu-name-confidence: http://www.gazbu.org/name-confidence/2.htm
latin-name: Krywy Taszleys
gazbu-latitude: 48.34947
gazbu-longitude: 30.89544
gazbu-location-confidence: http://www.gazbu.org/location-confidence/1.htm
geonames-id: http://www.geonames.org/830038/
contributor-id: mhp.techforge@gmail.com

Figure 5: Web page for *gazbu-entry* field.

Gazetteer of Beauplan's Ukraine
gazbu Database Field
.csv file field: gazbu-entry
.shp file field: gazbu-entr
Type: url
Description: Unique ID assigned to each entry in the gazbu database
Example: http://gazbu.org/entry/24.htm

Figure 6: Web page for *dgu* map.

Gazetteer of Beauplan's Ukraine
gazbu Maps
map code: dgu
map date: 1648
map name: Delineatio generalis Camporum Desertorum vulgo Ukraina
map source: http://catalogue.bnf.fr/ark:/12148/btv1b53041146h/

Figure 7: Web page for class = *riv*.

Gazetteer of Beauplan's Ukraine
Class: riv
Description: River
Location: Mouth of river
Associated gazbu field: gazbu-class

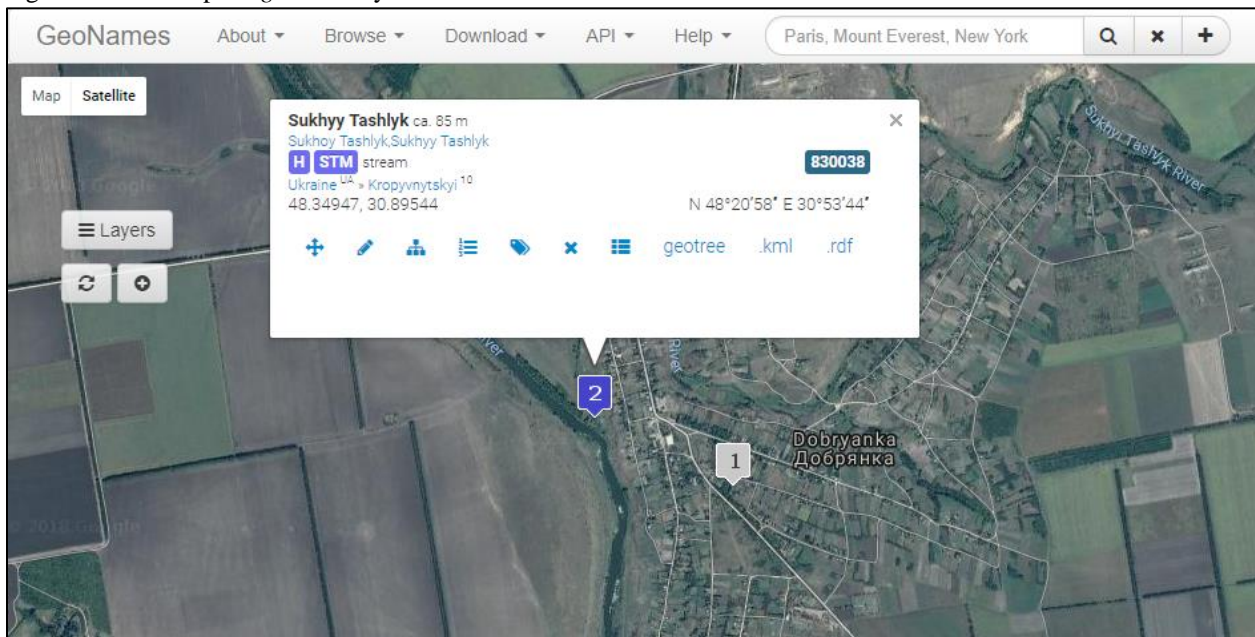Figure 8: Web map for *gazbu* entry 12.

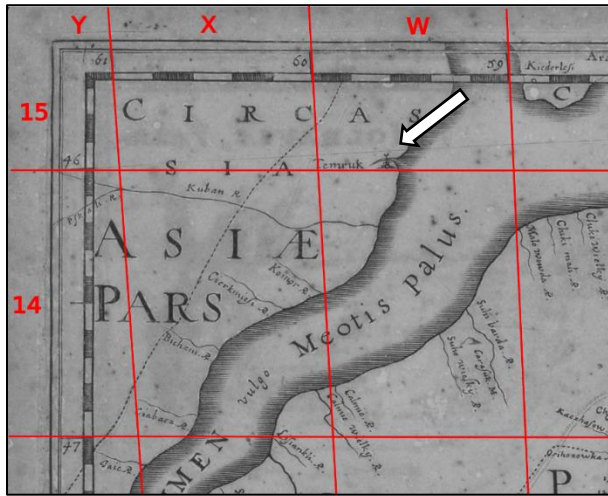Figure 9: Grid squares on *dgu* map (north at bottom).

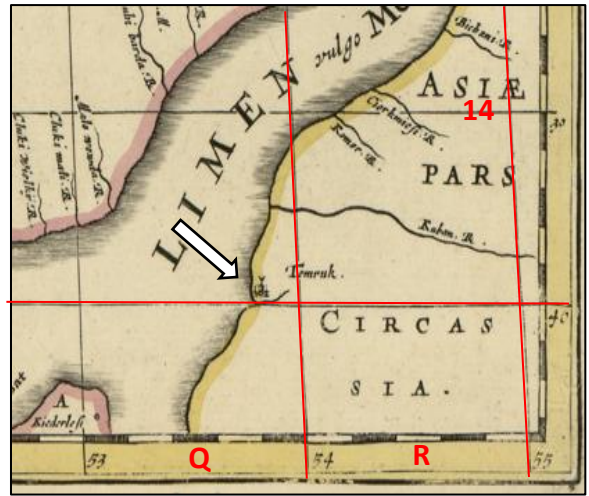Figure 10: Grid squares on *tgu* map (north at top).

Figure 11: QGIS project for creating *gazbu* database.

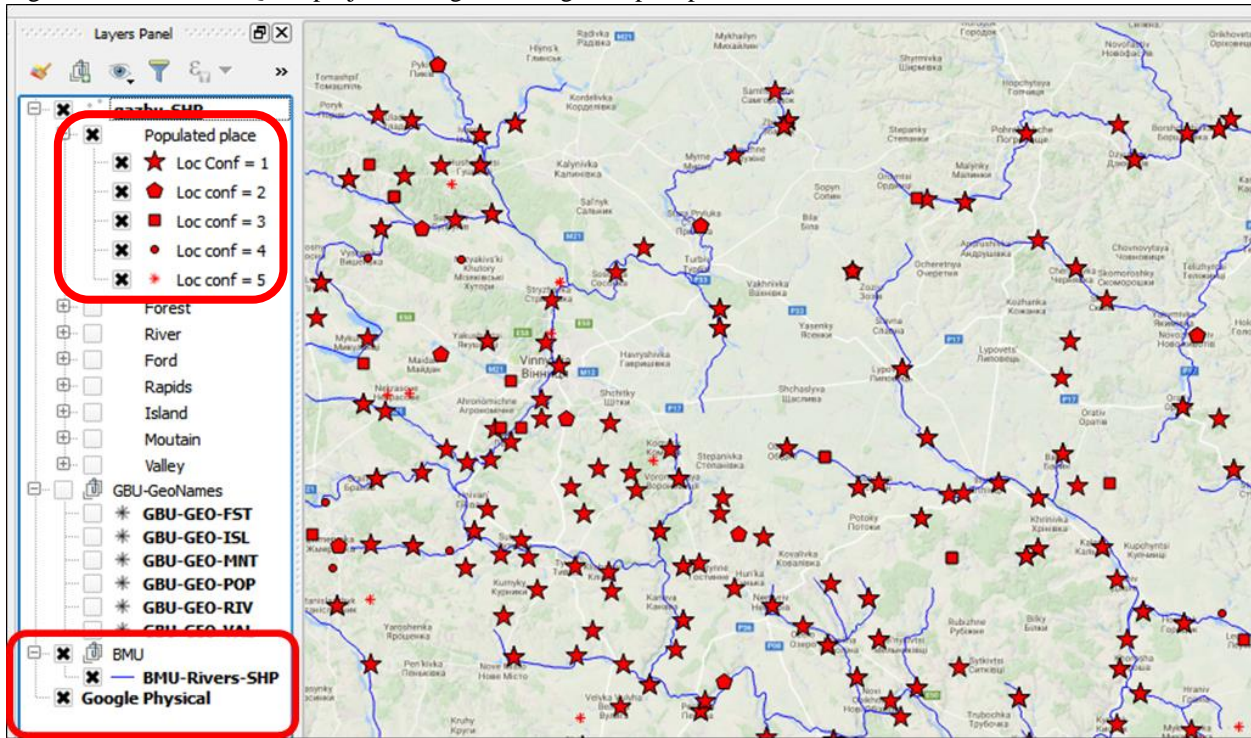Figure 12: Section of QGIS project after geolocating Beauplan places.

Figure 13: Beauplan map showing Czarny szlak.

Figure 14: Equivalent section of modern map.

Figure 15: Section of Beauplan map showing region of disputed territory.

Figure 16: Equivalent section of modern map showing region of disputed territory.