A new interactive gazetteer of Perczel's globe

Zsuzsanna Ungvári [a], Gábor Gercsák [a], Mátyás Márton [a], László Zentai [a,*]

[a] *ELTE Eötvös Loránd University, Institute of Cartography and Geoinformatics, Budapest, Hungary, 1st Author - ungvarizs@map.elte.hu, 2nd Author - gercsak@map.elte.hu, 3rd Author- matyi@map.elte.hu, 4th Author - laszlo.zentai@elte.hu*

* Corresponding author

Abstract: As a result of the Perczel Project outlined by Mátyás Márton in 2007 at the Department of Cartography and Geoinformatics of Eötvös Loránd University (ELTE), the complete digital restoration and then reconstruction of the 127.5 cm diameter manuscript globe was completed in 2019. The final cartographic completion of the ongoing task at the department by 2013 – led by Mátyás Márton, the project manager, who carried out this work with the help of Judit Paksi –, which included the work of many lecturers and students, aimed at saving the globe.

Using this augmented digital processing, it was also possible to create three artistic copies of Perczel's globe, which was originally made in 1862. This unique work of art, which has suffered irreparable damage due to the ordeals of the 20th century, is of great cultural value and also very important for our cartographic heritage, has been reborn.

Following the digital reconstruction, there was a good opportunity to revive another related project. The website, "The interactive gazetteer of Perczel's globe", which was born as a result of the work of Zsuzsanna Ungvári and Tibor Tokai earlier, was also created with a new, expanded content. The present study presents the antecedents of gazetteer creation and describes the most important steps of current processing.

Keywords: virtual globe, interactive gazetteer, name geodatabase, name reconstruction and restoration

1. Introduction

During the digital reconstruction of Perczel's globe, the authors met several challenges. The most time-consuming and largest job was the reconstruction of the names. At the beginning of the project, the idea emerged to create an interactive gazetteer of this giant globe made in 1862.

The digital restoration for this part of the project meant the identification of legible names character by character. The illegible characters were marked by question marks, if it was possible to count or estimate the number of missing characters. Unfortunately, a significant part of the names is badly damaged by now, but the lacquer layer, which coated the globe surface from the 1970s, turned yellow-brown soon and dissolved some parts of the ink.

In the first phase of the restoration, it was not allowed to find out the illegible characters: the only way to complete these letters was if the students had found contemporary maps that proved their idea for the examined name.

Since it needed too much time and energy to read the poorly visible letters, lots of names contained question marks. It was not rare that the whole name was replaced by question marks only (Figure 1).

At the end of the work, the authors created a digital facsimile on which the graphic content and the names have saved the original state of Perczel's globe. The font families imitated the handwriting. The colours of the geographic features followed the original ones, which is known from an earlier description (Ambrus-Fallenbüchl, 1963).

Some parts of the globe map are badly damaged or came off from the globe surface: these parts are not or partially patched in the digital facsimile. Most part of the restoration of the globe was completed in 2012 (Figure 2).
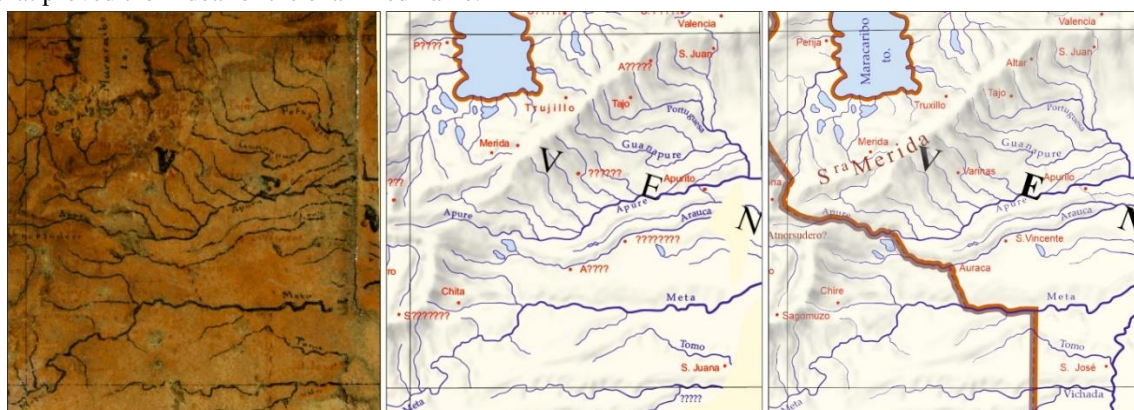

Figure 1. Digital restoration (2012) and reconstruction (2019) of geographic names

The real, physical, artistic facsimile of the globe was completed in 2020, but the restoration was continued in the previous year too. The tasks included correcting the map, completing the hill shading and the names in the torn off areas. The greatest challenge was to find the correct contemporary name forms: after carefully examining the surrounding labels, it was possible to determine the source maps that László Perczel possibly used. We had to take into consideration that Perczel mostly preferred the Hungarian form of names. Unfortunately, Perczel did not use these Hungarian form of names consequently. This statement is especially true for the appellatives. Here you can read some examples: the names in the first column were used by Perczel, the second one includes the names as written in a possible German source atlas (Universal-Handatlas), and the last column gives the names as they are in their today's Hungarian form.

Perczel's form	Universal Handatlas	Today's form
Kanin fsz	Kanin	Kanin/Kanyin-fsz.
Archangel	Archangel	Arhangelszk
StPETERVÁR	St. PETERSBURG	SZANKT-PETYERBURG
Moskva	Moskau	Moszkva
Elborrs hegység	Elburs Gb.	Elburz
Persiai Öböl	Persischer Golf	Perzsa-öböl
Ormus Szorasut.	Str. v. Ormus	Hormuzi-szoros
Jukon	Jukon	Yukon
Yellov Stone	Yellow Stone	Yellowstone
Misisipi	Mississippi	Mississippi
Hudson utja	Hudson's Strasse	Hudson-szoros
BUENOS AYROS	Buenos Ayros	Buenos Ayres or Buenos Aires
RIO JANEIRO	Rio de Janeiro	Rio de Janeiro

Table 1. Examples of the globe's names in a contemporary source and their today's form.

The Universal-Handatlas was edited by Heinrich Berghaus in Glogau (today Głogów, Poland), 1859. Perczel may have used this atlas while editing his globe: this is proved by numerous graphic and onomastic similarities. He kept the German form of names in several cases. This atlas had earlier publications: for instance, the 5th edition in 1857 or earlier could be his source.

2. Antecedents of the interactive gazetteer

2.1 The first version

The first collection of the names was carried out by three BSc students in undergraduate research and BSc thesis in the same year, which contained European, African and South American settlement names (Nyuli, 2009; Tóth, 2009; Való, 2009; Tóth et al., 2009). The structure of the database were a little bit different from the recent one (Table 2). The first column contains the names that they could read (that were clearly legible) on the globe map, and the second column shows the recent name of towns.

Read name	Probable name	Legible?
Tananarivo	Tananarivo	True
M??umba	Mayumba	False
Malenba?		False

Table 2. Examples from the first version of the dataset

In addition, the geographical coordinates were saved in the database. The final result was published in a KML file.

Two years later, another BSc student created a web-based viewer application to this geodatabase in his BSc thesis (Tokai, 2011). The KML file was imported to a MySQL database. The website was connected with this collection, in which users were able to find the names and viewed them on a virtual globe provided by the Google Earth (GE) plugin. The GE globe was covered by a raster map of the original Perczel's globe. Ten years ago, the usage of this plugin in the browser was an innovative and almost the only available solution to make an interactive globe gazetteer on the web.
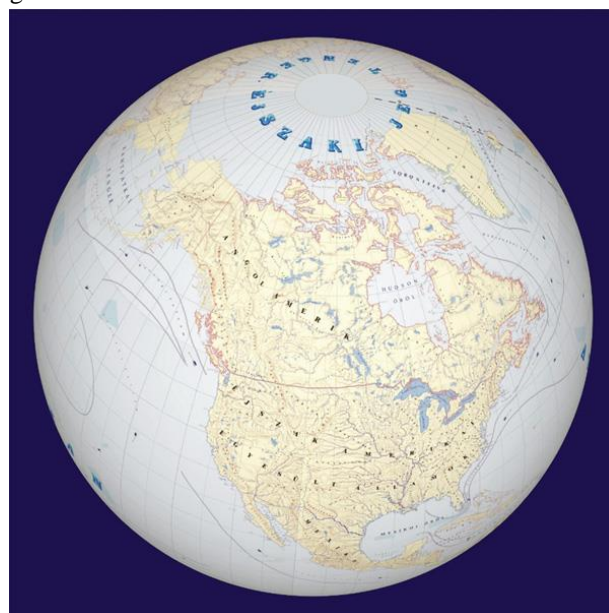

Figure 2. The virtual facsimile (2012)

2.2 The second, extended version of the gazetteer

In the next years, the first author of this paper took over the organization and the technical management of the interactive gazetteer project, and published papers on it in the national (Ungvári, 2012) and international scientific scene (Ungvári et al., 2013; Ungvári, 2014).

Several excellent students took part in the name collection process: they identified not only the settlement, but also other geographic names (e.g. rivers, countries, mountains). The new database got an updated website, which was based on the first version, but extended by several new functions (Figure 3), these were:

- the geographic names and the settlement names were put into two (separated) list because of the large number the city names, and the different

view of the storing of coordinates (explained later),

- the possibility of filtering names by continents,
- the new, reconstructed globe texture was available beside the original,
- setting the transparency of the globe texture with a slider,
- the names in KML files were downloadable.

As said, the settlement names were recorded by their geographical coordinates in the database, while the other geographic names were added by a code, which came from a search grid. In a gazetteer to a map, when the users search for names, they find the geographical objects by grids. Therefore, for easing the search in grids, the globe's geographical grid by 5 degrees was chosen. Each square was numbered from 1 to 72 by five degrees between the longitudes; the squares between latitudes were lettered A, B, C, etc. by five degrees moving away from the Equator to the poles. To avoid the same lettering on the globe, the squares on the northern hemisphere got an "N" prefix, and an "S" prefix on the southern one (e.g. NA, SA).

The advantage of this method is the faster registering of the elements in the geodatabase

However, in practice, this worked in opposite way in the website. The code of names with compact extension like river or volcano names were simple, but the stretched linetext objects like Orosz Birodalom (Russian Empire) or Sahara Sivatag (Sahara Desert) needed the code of two geographic rectangles. The website zoomed into the rectangle(s), and the users had to find the names somewhere inside the rectangle, but it was not easy where the tint was partially or fully illegible.

2.3 Shut-down of GE plugin

Unfortunately, Google shut down the Earth plugin in the beginning of 2017, so the new users were not able to install the plugin. Therefore, the viewer surface of the website had to be renewed.

Even in this year, the author updated the 3D virtual globe to Cesium JS, but no other changes were made.

Figure 3. The earlier version of the interactive gazetteer website in Hungarian

3. Revitalization of the Perczel Project and the interactive gazetteer

As Márton et al. summarized in the article (Márton et al., 2021), the order of making three real reconstructed copies of Perczel's globe gave new dynamism to the project: the names on the whole globe were examined and completed. Contemporary source maps and atlases were used to identify the corrupted names.

After the birth of the three real globes, the first author began to refresh and rebuild the geodatabase.

The virtual reconstruction included two times 36 pieces of 10-degree wide maps in Cassini-Soldner projection from the Equator to poles. The georeferenced segments were assembled in plate carrée projection (Gede, 2009), and this map was imported to GIS software (QGIS).

In the case of settlement names, the first author only refreshed the names and the position of symbols. In the case of other geographic names (rivers, lakes, mountains, etc.) the 5°×5° grid rectangles were dropped. The names with shorter extension like river names, volcanos, smaller bays are recorded by their horizontal coordinates.

The geographic names with significant extension are indicated by a path (polyline) or polygon. This made the processing of data somewhat complicated, but we get a much more precise location of names when the users zoom onto the selected object.

Figure 4. The renewed website in Hungarian

The question can come up, "Why do not we store the whole object as points/polylines or polygons in the database? And why do we not show these geographic objects in the viewer window?" The answer is short: this eventuates in travelling too much data (almost 8,700 objects!) through the network, which makes the browsers slow or get them frozen. The Cesium JS cannot work with so many features. Transferring the raster texture in appropriate resolution can be slow if the network conditions are poor. An alternative solution can only be loading the searched object in vector format, but the symbolization of vector data in Cesium is not so professional as we need it. Only the usage of raster texture can give back the real value and beauty of this globe.

Not only the database, but also the website has improved. Beside filtering by continents, it is possible to reduce the results by the type of geographic objects. The two separated lists for settlements and other geographic names are preserved. The technological background has not been changed since the last update. A short description of this project in English and Hungarian has been added to the website.

4. Access to the interactive gazetteer

The project website is available at http://terkeptar.elte.hu/vgm/perczel. The readers may also visit the Virtual Globes Museum homepage first, where the "Related Projects" offers the Interactive gazetteer of a 150-year old globe (Figure 4).

5. Statistics tells more

To feel the size of this project and the difficulty of identifying names, some statistical figures on names tell the reader much more. The total number of names on Perczel's globe is 8,874. The globe's diameter is 127.5 cm, which corresponds to a 1 : 10,000,000 scale map. Table 3 summarizes the type of the names in the final database.

Name type	1	2	3
Settlements	4,258	1,408	33.1%
Islands or archipelagos	1,076	23	2.1%
Reefs	24	0	0.0%
Rivers	1,727	89	5.2%
Lakes (incl. moorlands)	219	1	0.5%
Seas and oceans	44	0	0.0%
Bays	183	3	1.6%
Capes	339	13	3.8%
Mountains	269	27	10.0%
Volcanos	110	11	10.0%
Straits, passages	69	0	0.0%
Ocean currents	27	1	3.7%
Nations	45	1	2.2%
Administrative names (country, state)	232	6	2.6%
Other areas (incl. deserts)	226	15	6.6%
Other point features	20	4	20.0%
Grid names	6	0	0.0%
SUM	**8,874**	**1,602**	**18.1%**

Table 3. 1: The number of features in the category. 2: Number of features in the category whose names contain minimum one question mark. 3: Percentage of the name with question mark.

The result is not accidental. The reason of the high number of illegible settlement names (written in red tint) is the use of the yellow/brown lacquer layer that covers the globe surface. This cover was added during a poor-quality physical restoration in the 1970s. The red ink has been damaged and dissolved by the lacquer. Although the font size of river names is almost same as that of settlements, but they were written in dark blue ink, which explains why the rate of illegible river names is much lower.

The authors also collected those settlement names which were only marked by three question marks, because they were unknown names: this meant 144 items (3.4%). All

other names contained minimum one letter that helped their identification.

To illustrate the changes in the database against the earlier version, the authors made some queries. In the last version of the database, 7,302 names were introduced (3,403 settlement names, and 3,899 other geographic names). This number is 1,572 less than the final one. This difference comes from two sources: new names on the supplemented areas and on the areas that were not processed in the former version (e.g. Asia Minor, parts of Australia).

In the former dataset, the number of names with one question mark was 2,580, from which 2,322 (this is 68% of 3,403 settlement names!) were settlement names and 258 other geographic names. If you compare it to the recent value (33%), the author reached much higher rate of identifying city names. Totally 3,252 name were supplemented or replaced on the globe.

6. Outlook

This interactive gazetteer is based on globe visualization, but this is not unique in the web cartography: the undersea feature name gazetteer created by GEBCO also uses the Cesium JS library to represent the geographic object's recent name (GEBCO website 2021).

7. Summary

The authors outlined the challenges and the final result of the interactive gazetteer of the Perczel's globe project. The name dataset that the contributors had collected in the first years has been completed and totally revised. Numerous corrections of names were introduced into the database. Beside this, the webpage of the project was renewed.

8. References

Ambrus-Fallenbüchl, Z. (1963). Der grösste Erdglobus Ungarn – hundert Jahre alt. Der Globusfreund, Publ. Nr. 12, Wien

GEBCO website (2021). https://www.gebco.net/

Gede, M. (2009). The Projection Aspects of Digitising Globes. In: Hunt E. et al. (ed.): Proceedings of the XXIV International Cartographic Conference. Santiago, Chile, 2009. Paper 10/1.

Márton, M., Gercsák, G., Zentai, L. (2021). The Perczel Project (2007–2019). In print.

Nyuli, É. (2009). Európa településeinek térinformatikai adatbázisa a Perczel gömb alapján, *A spatial information database of European settlements on the basis of Perczel's globe.* BSc thesis, ELTE Eötvös Loránd University, Budapest, 2009.

Tokai, T. (2011). Webes interaktív névmutató készítése virtuális glóbuszokon. *Making an interactive web index of names on virtual globes.* BSc thesis, ELTE Eötvös Loránd University, Budapest, 2011.

Tóth, B. (2009). Dél-Amerika településneveinek geoadatbázisa a Perczel-glóbusz alapján. *The geo-database of the settlements of South America on the basis of Perczel's globe.* BSc thesis, ELTE Eötvös Loránd University, Budapest, 2009.

Tóth, B., Való, A. (2009). Geoinformatika a kulturálisérték-mentés szolgálatában (Régi térképek földrajzi neveinek adatbázis-építése). *Geoinformatics in the service of saving cultural heritage (Building databases of names of old maps).* Undergraduate research at the Hungarian Undergraduate Research Competiton. ELTE Eötvös Loránd University, Budapest, 2009.

Ungvári, Zs. (2012). Névrajzi keresőrendszer kialakítása a Perczel-glóbuszokhoz. *Developing a search system of names for Perczel's globe.* Presentation in Hungarian at the Symposium organized for the 150-year-old birthday of the globe. http://lazarus.elte.hu/hun/buszke/2012-perczel/perczel.htm

Ungvári, Zs., Tokai T. (2013). The interactive gazetteer of a 150-year-old-globe. In: Manfred, F. Buchroithner (ed.): Proceedings of the 26th International Cartographic Conference. Dresden, Germany: International Cartographic Association, (2013) pp. 721–722.

Ungvári, Zs. (2014). A Method to Create Interactive Gazetteer to Old Globes In: Livieratos, E; Pazarli, M. (eds.): 9th International Workshop on Digital Approaches to Cartographic Heritage. pp. 280–287.

Való, A. (2009). Afrika tepülésneveinek geoadatbázisa a Perczel-glóbusz alapján. *The geo-database of the settlements of Africa on the basis of Perczel's globe.* BSc thesis, ELTE Eötvös Loránd University, Budapest, 2009.