Andrea Pődör*

Experiment of Involving Students in Preserving Geographical Names Appearing on historical maps

Keywords: historical geographical names; crowdsourcing; data acquisition

Summary: Due to historical changes many toponyms may disappear or change throughout the centuries, even shorter periods can cause relevant transformations. Although these knowledge would explain many historical facts in the past, these data inevitably should be part of national data stores. Recently, many attempts show that crowdsourcing or citizens' involvement can be used for data acquisition, therefore the aim of our study is to investigate the possibilities in crowdsourcing in collecting toponyms from historical maps. In last year's fall semester 18 students of the University of West Hungary participated in a preliminary experiment of the National Széchényi Library and the University of West Hungary. In the experiment we analyzed the effectiveness and accuracy of the students in collecting geographical names during a certain amount of time. Our results prove that involvement of students/citizens can be a solution but preliminary education and background knowledge would be necessary for a successful project.

Introduction

Geographical names have an impact on many areas of life like economy, science, business, planning etc. Geographical names also serve an essential part of national treasures, each nation's duty is to preserve as much geographical names, as possible. This knowledge can support further historical and geographical research and also can explain historical and geographical events in the past. Additionally, these names can change and disappear throughout the centuries. From this point of view Hungary has a special situation because after World War I a considerable amount of its former territory belongs to other countries where the official languages are different, therefore the original Hungarian geographical names are disappearing. "Perhaps all Hungarian geographical names originally had an informative meaning for the Hungarians or other people, but due to the course of history and migrations many of them became meaningless for the present population" (Gercsák, G. 2001). Naturally, in many cases their original meaning can be felt or detected. The experts of National Széchényi Library take notice on the importance of old toponyms, they initiated many project for the preservation of geographical names. As toponyms part of the national treasure and in the case of Hungary it is specifically important that these names are widely available for the public, because most of the names referring to places outside of the recent boundaries. (Ungváry, R.& Pászti, L. 2005, Ungváry, R. 2013).

Preserving toponyms is especially time and human resource intensive, so implementation of crowdsourcing seems to be a promising solution to make use of. On the other hand the special grammatical rules behind geographical names in Hungary make it a more complex task to deal with by a non-professional citizen, so involving students in geography, cartography or surveying may be a good answer for this problem.

* University of West Hungary, Faculty of Geoinformatics [pa@geo.info.hu]

Databases in Hungary for geographical names

In Hungary there are three main namespaces available:
1. The Gazetteer of Hungary (GH) elaborated in the 1960's is the main source of geographical names for map design (Földrajzinév-tár, 1984.) The Committee of Hungarian Geographical Names (Magyar Földrajzinév-bizottság) has authority on editing and approving the content of GH. This database is treated by the Institute of Geodesy, Cartography and Remote Sensing (Földmérési és Távérzékelési Intézet, FÖMI), data records of GH can be purchased in table or text form, but there are approximately 79000 records with missing geographical location. (Guszlev et al., 2007).
2. KÖZTAURUSZ and Geotaurusz is the thesaurus of geographical concepts is maintained by the Hungarian National Library. It contains 73 000 records, it can be accessed freely, and it is managed by the RELEX software. The RELEX can manage the integration of new information, such as added notes, relationships and new terms and UDC codes. It also provides the export of records according to the xml-based OWL Web Ontology Language (Ungváry et al., 2012). This thesaurus is continuously updated and enriched by the remarks of librarian, with analysing maps and using other databases like GH.
3. The Hungarian Digital Toponym Registry (http://mdh.unideb.hu/en/) is managed by the University of Debrecen. It contains 896248 contemporary records and 17986 historical ones.
In the outlined experiment the author used the RELEX software.

Experiment

For the experiment the experts from National Library of Hungary chose an administrative map of the Hungarian Kingdom which was created before 1910 (Figure 1). In 1910 all the Hungarian toponyms were regularized in order to avoid homonyms which were quite frequently used at that time. After 1910 a lot of toponyms were officially changed according to the rules.
The above map was compiled in 1884, its scale is 1:360000. The map consists of 12 tiles, and besides placenames it contains the names for mountains and watercourses.
The typography of the names in the map indicates the type of the geographical name. The following types can be distinguished:

- Semibold normal: Placenames,
- Semibold normal with capital letters: Cities
- Semibold italic: mountains, ghaut
- Italic: villages
- Calligraphic style: part of a place, (many of them became independent settlements recently)
- Italic in blue colour: watercourses.

In the fall semester 18 students of the former University of West Hungary Faculty of Geoinformatics participated in the experiment to test Project based learning on geographical names and RELEX database.
The age of the students were between 19 and 25 years, four of them studied in the second, the rest in the sixth (last) semester. Only four of them were female.

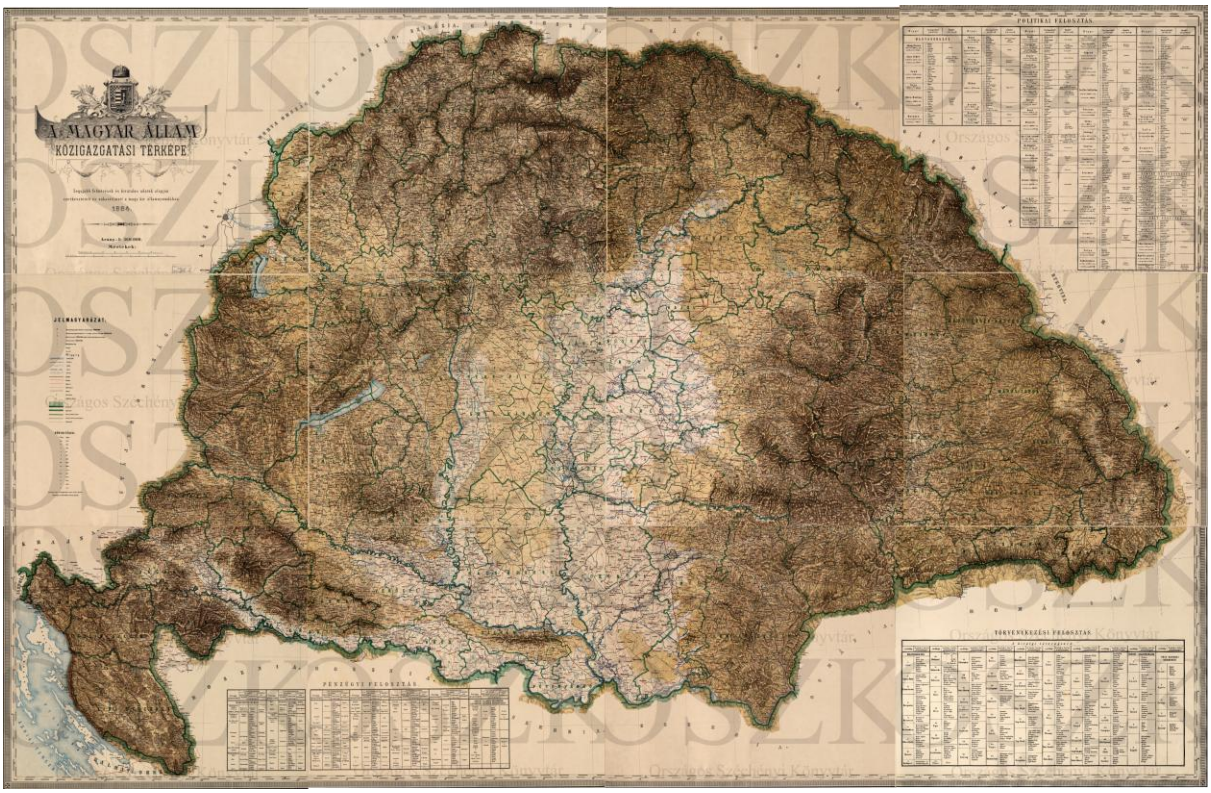

Figure 1. The administrative map of Hungary which was used in the experiment.

The students are studying to be a chartered surveyor, they have a course for cartography for one semester. The topics of the course cover all the main issues from history of cartography to thematic mapping and geovisualization. They have several courses on GIS and related field as well.
In the experiment the author wanted to test the effectiveness of the involvement of students in gathering and building a repository of historical geonames as a part of the national heritage.
The author had other experiences based on a type of project-based learning (Pődör, 2011), where the students started to work together and later formed small groups, where they worked on their own. In this experiment the author followed a similar sequence of activities.
Before the workshop they got some material for homework to familiarise themselves with the grammatical rules and forms of the Hungarian and foreign geographical names. The experiment started with an educational workshop about geographical names by experts from the National Széchényi Library (Rudolf Ungváry and Eszter Kiss). They started with an overview on geographical names, thesauruses and the basics of ontologies. This introduction lasted for two hours. Following that part the students formed two groups of 9-9 participants. The expert spent further two hours with each group to provide hands on experience with selected maps and the RELEX database. At the beginning Rudolf Ungváry, expert of RELEX database explained the idea, aims and basic of the usage of the database. He created a special training domain for the students, they have got a password and access to the database with the right to upload new data into the database. He also prepared a document in order to help the students with using the database. Other documents were available on the E-learning platform of the course together with the handbook of the database, and other sources like Google and MAPIRE.
During the practical part one tile of the map series was chosen, namely one map tile covering a part of Transylvania with the surroundings of Gheorgeni (Gyergyószentmiklós) and the group directed by the expert of the library followed the geographical names on this tile horizontally. Initially the experts showed the method the students had to follow, namely: 1) define the type of

the geographical names 2) if the name was difficult to read how to clarify with the help of Google and MAPIRE, 3) the last step was to check the status in the RELEX database.
Depending on the status of the name there were different tasks to do:

- the processed name already existed in the database exactly at the same form – in this case there was nothing more to do,
- the name was in a different form possibly because it was in an older form – in this case the students had to upload as a new name but with the relation of L (see below/vide infra),
- if the name did not exist in the thesaurus they had to upload as a new item but they had to define a larger geographic location:
    - for the periphery is the settlement,
    - for the peak or the mount is the mountain
    - for the settlement is the township
    - for the township, landscape, hill is the county
- in the case of water network with the relation (R), they had to define in which other river the given river fell into,
- if the origin and meaning of the name could not be clarified, they had to indicate it in the database
- if the name is a homonym, it means it is included in the thesaurus, but it depicts a different geographical location, so it should be uploaded and then the larger geographical location should be indicated (©Ungváry, R.).

In some cases the map was very difficult to read, therefore the students were asked to use two reference maps: Google, and MAPIRE which contains the three military mapping surveys of Austria-Hungary Empire, made before and during the first years of the First World War. (http://mapire.eu/hu/map/mkf_hun/?zoom=6&lat=47.11091&lon=20.54031) .

Results

In the week following the workshop the students had to continue the map analysis on their own without any external help, for one and half hour time. They could form a group of 2 students and they could share their work. One of them could handle the RELEX software and the reference maps while the other could analyse the processed map and document their findings in a word document. Also there were some students who found out that working alone is more effective for them so they analysed their map tile alone. As they were not familiar with names outside the recent boundaries of Hungary, most of them could choose a tile for analysing which is well-known for them, for example the area where they come from. During their work they processed names from the following counties of Hungary: Baranya, Bács-Kiskun, Fejér, Jász-Nagykun-Szolnok, Komárom-Esztergom, Pest, Somogy, Szabolcs-Szatmár-Bereg, Veszprém and Sălaj (Szilágy in Hungarian) county from Romania. All of the students processed 259 names altogether of which 52 are new geographical names (see Fig. 2).

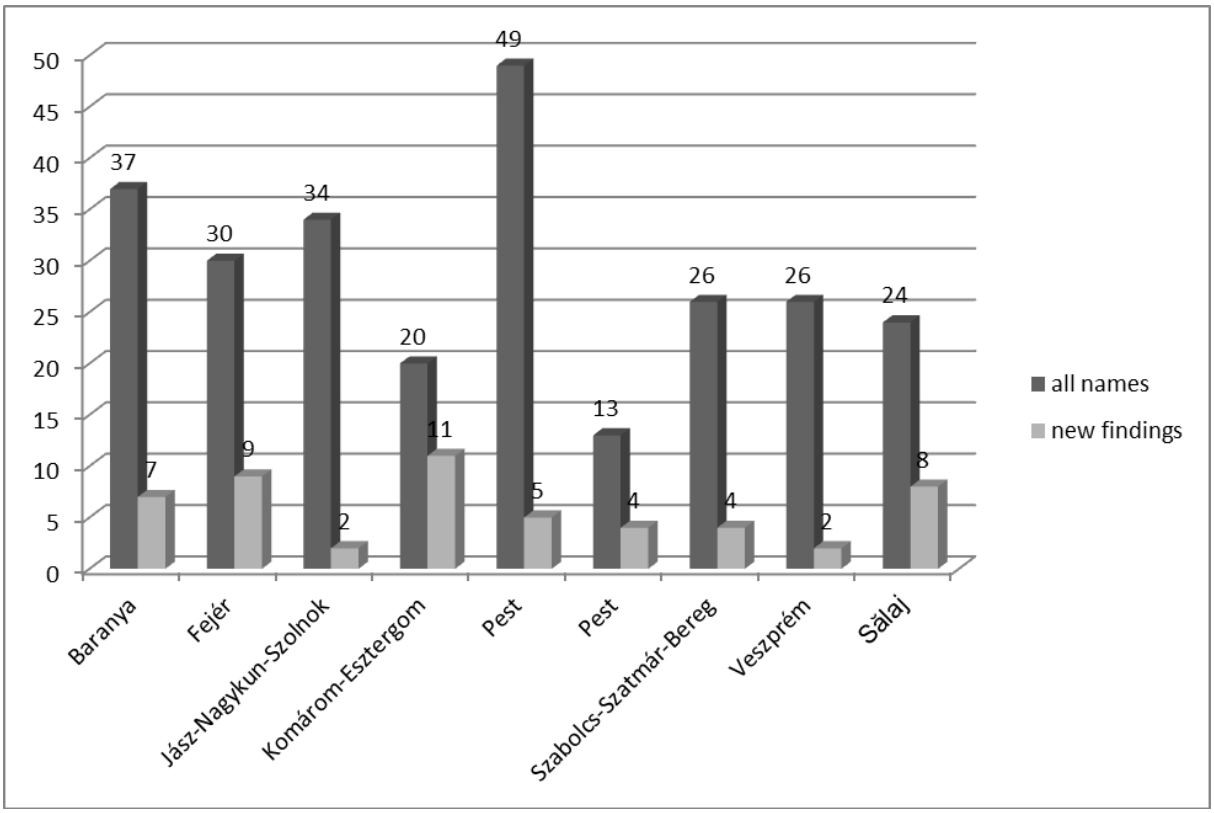

Figure 2. The results of the experiment.

The student, who processed the most names (49) worked on Pest county, although he found only 5 new ones to upload into the database. Those who found relatively much more new items were working on Komárom–Eszetrgom and Sălaj county. In case of Komárom-Esztergom most of the findings were part of settlements (8). In case of Sălaj the reason is probably that these names are now outside of the recent area of Hungary, therefore this part is not so well documented in the existing database. The students could process approximately 28 names on average including each type of geographical names. It was sometimes quite time consuming to check the names in other sources, and the map was not easy to read.

If we control visually the improvement of the students (Fig. 3) it is obvious that areas which are quite flat and not covered with drawing of slopes are much more easily readable for the students. Although it is interesting that they mainly processed place names, there were several other type of names in the maps. Only one group of students were dealing with names of watercourses. It is also interesting how they covered the area when investigating the names. As we can see in Figure 3, these students mainly worked along the Danube river and not horizontally on the map segment. It can be also seen that they omitted some names, probably they were not able to read or identify them.

Generally we can state that the students tend to work very accurately, although it is possible that during the uploading process they made mistakes identifying the relations, so the uploaded names should be supervised by experts from the National Library, as well.

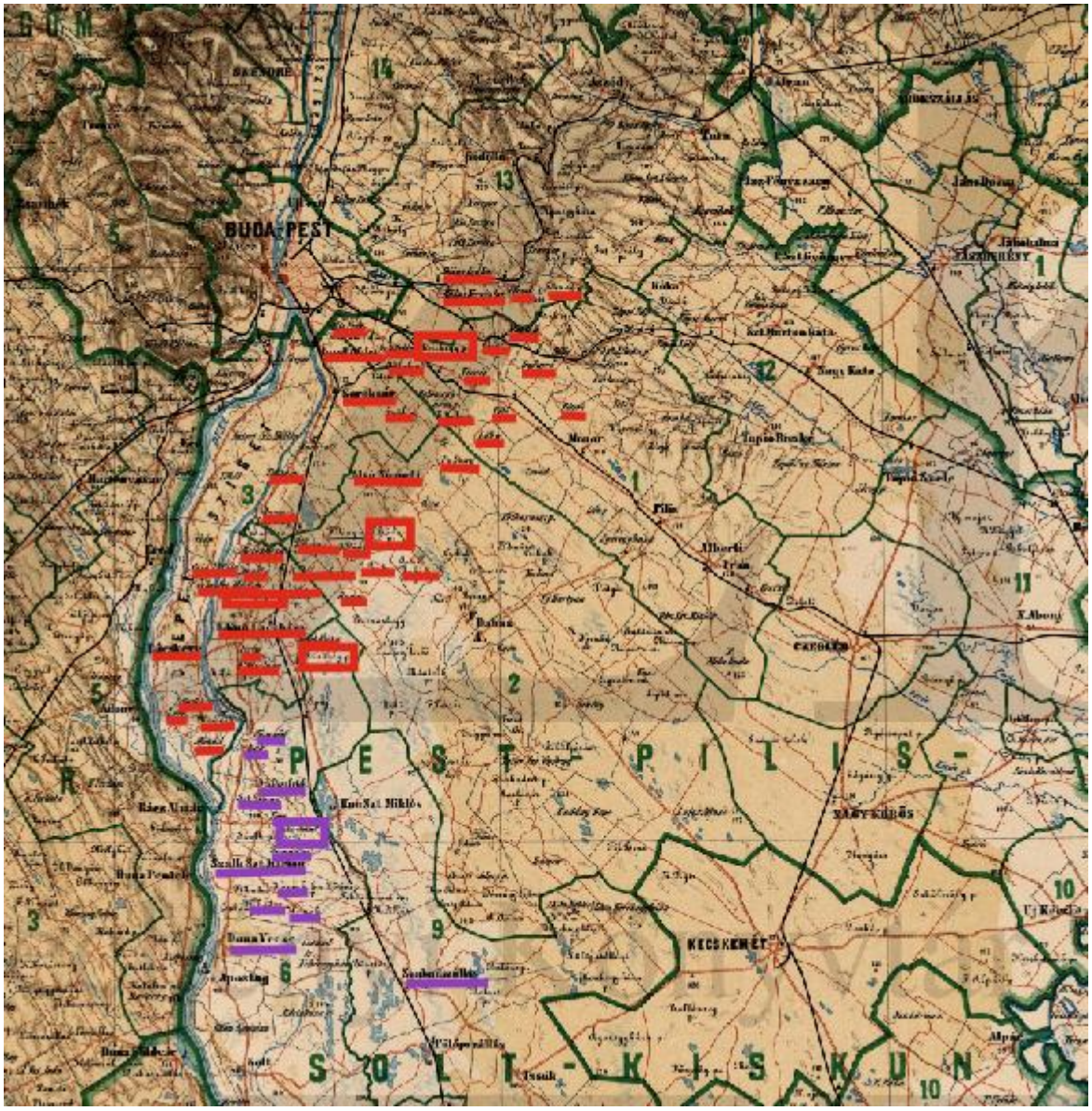

Figure 3. The visual interpretation of the work of 2 students.

Conclusion

The collection of geographical names is very important as these are part of the national history and treasures. In our experiment we tested the effectiveness and accuracy of collecting geographical names by students during a certain period of time. Our results prove that involvement of students/citizens can be a solution but preliminary education and background knowledge is necessary for a successful project. We also realised that this process is very time consuming and the involvement of more students is desirable. Also the author thinks that gamification can enhance the students' productivity as well. As a foundation of the further geographic name collection projects the cooperation between the National Library and Universities is necessary, although the collection process should be more elaborated to make sure that all the names in the map analysed will be included.

Acknowledgements

The author would like to thank Rudolf Ungváry for his valuable comments and help in the workshop, and András Révész for refining the English text.

Literature

Gercsák, G.: "Using Hungarian geographical names in English language publications". In: Studia Cartologica Vol. 12, Eötvös Kiadó, Budapest, 2001, pp. 171-180.

Guszlev A, Lukács L, Reyes J: Database of Hungarian Geographical names. In: Alexandr Borodko (eds.) Cartography for every one and you: Proceedings of the 23th International Cartographic Conference, 2007.08.04-2007.08.10, Paper 3-3-2.

Pődör, A.: The Methodological Advantages of Using Web Server in Teaching GIS. In: Ruas Anne (ed.), Advances in Cartography and GIScience, Volume2. Heidelberg; New York: Springer Verlag, 2011, pp. 73-82.

Ungváry R. -Király L. -Szabó Gy.: The Relex of the Web. A freely usable for construction and maintenance of name space, thesauri and ontology. Networkshop Conference. Veszprém. 2012 http://nws.niif.hu/ncd2012/docs/aen/031.htm

Ungváry, R.: A földrajzi névtér jellemzői. Networkshop Conference, Sopron, 2013. nws.niif.hu/ncd2013/docs/phu/069.pdf

Ungváry, R.-Pászti, L.: A földrajzi nevek mint az adatbázisrekordok hozzáférési pontjai. In: Könyvtári Figyelő, 51. köt. 2005. 4. http://epa.oszk.hu/00100/00143/00057/ungvary.html