SECTION 4. Designing and testing online database for toponyms

A special concern for toponymy in a project which is essentially one of „landscape archaeology", is rather unusual. The situation owes to some characteristics of the northern area of the track (north of Urlueni) which are hard to explain if we are to judge them according to the landscape we see today or the one that can be reconstructed according to old maps; the oldest maps with details about vegetation (so at a better scale than 1:50.000) doesn't go any further than 1790 (Specht map), in other words in modern times (according to Occidental criteria, more and more accepted in our space too). At the end of the 18th century, the Romanian Plain had already passed through 4-5 centuries of colonization and agriculture, thus changing its original landscape.

The idea of a detailed study of tononymy – as adjuvant for the cartographic study – owes to the simple observation that the names are older than their attestation (on a written or cartographic document). When the toponyms doesn't invoke persons, they evoke natural elements that may be a substitute for the absent descriptions of physical geography. For instance, only during this incipient study there were found tens on names that suggest forests, in areas in which no forests are documented.

Figure 4.1. The main form of the database for toponymy.

According to scheduling, a database was designed for toponyms, using the support offered by Microsoft SQL, a language that allows the use of data in applications like MS Access, but also in web interfaces. The most important aspect is that even in this stage the partners can follow each other's progresses, the database being accessible via Internet.

We chose a simple and flexible structure of data. It includes (v. fig. 4.1.) the following:

- name
- primary classification of toponyms
- superior administrative unit
- SIRUTA code
- hydrographic basin
- coordinates in Stereo70 (official projection in Romania)
- geographic coordinates (useful for non-specialists, for instance in Google Earth)
- document data
- signature
- date
- research project (in this case *Limes Transalutanus*)

Several explanations should be given about the aforementioned fields:

- the name is transcribed as close as possible to the document's original orthography, but it is doubled by a transcription in today's phonetics (for instance „Ciochesci /= Cioceşti/"), in order to facilitate the identification of all the names (or orthographic variations) that a place had;
- the toponyms' taxonomy has, at this moment, the following structure:
    - confluences[1];
    - hydronyms (which doesn't refer only to streams, but to any element that invokes water, like fountains or everglades);
    - settlements – which are, at this moment, all the inhabiting resorts, like monasteries or rest houses (inns);
    - mounds (include hillocks, due to a language particularity within the region);
    - major landforms;
    - forests (with all their nuances, up to bushes);
    - toponyms (that cannot be included in any of the classes above)
    - finally, explicit anthroponyms will be added to the list above, like the names of the persons who are mentioned in the medieval documents; the reason is simply that we are trying to locate some families, areas, estates, on the chance that the toponyms list and the anthroponyms list will find some common terms.
- the superior administrative unit refers to county, but adapted to specific situations in different moments in time, of course on the condition of knowing the form of administrative organization from that specific time;
- SIRUTA code is the unique code for identification of a locality or administrative entity (parish or county), and will be used with the purpose of grouping all the types of toponyms existing *today* within the perimeter of an entity;
    - It would be ideal to use the codes of settlements, which are the most stable territorial unities (despite the frequent changes of name); alas, the limits of villages within town hall's area is not a public resource at the moment;
    - in this situation, the SIRUTA codes of the parishes will be used, despite the frequent changes of their composition;

---

[1] On the strength of the observation that most of the settlements (especially the large ones) are not only located „on a stream", but at the confluence of running waters, thus offering to the community several water sources, positioned in areas of the commonly used territory (for agriculture or livestock breeding).

- the hydrographic basin is a principle of segmenting the territory according to criteria of connectivity, of gravity around some vital life resources (here: water); as „hydrographic basin" were selected streams and their confluents in which at the present there is a group of minimum 6 villages; the hydrographic basins were obtained through semi-automate procedures, that is by getting automatically the valley floors from a SRTM, while the frontier was established on watersheds, dependent on the presence of a certain number of villages;
    - 24 hydrographic basins were defined, from west to east: Olt, Dârjov, Iminog, Dunărea, Călmăţui, Călmăţuiul Sec, Urlui, Bratcov, Dorofei, Plapcea, Negrişoara, Vedea, Vediţa, Cotmeana, Tecuci, Burdea, Tinoasa, Pârâul Câinelui, Teleorman, Bascov, Argeş, Albota, Dâmbovnic, Neajlov;
    - there are also territorial limits for the research at the project *Limes Transalutanus*, which are approximately 20 km west and east of the old Roman frontier, so the aforementioned hydrographic basins are analysed only insomuch as they overlay the limits of interest;
    - the small hydrographic basins have from 66 km^2 (like Negrişoara) to 166 km^2 (like Dârjov), while the large basins, like Vedea, reach up to 400 km^2 (although incomplete);
- the coordinates – be they Stereo70 or WGS84 – were copied from the properties of the point, as they are displayed in the GIS file (see below); they are not destined only to facilitate the identification of a toponym, but also to allow calculations of distances and directions between two specific points;
- document data refers either to the publishing date of a map (or map sheet), or to the issuance date of a document (medieval, for instance);
- signature (by acronym) of the person who delivers the data;
- date – is completed automatically by the database;
- research project – field that has been introduced on the chance that other volunteers or groups of researchers will contribute to the database, for other areal than the ones defined for *Limes Transalutanus*[2].
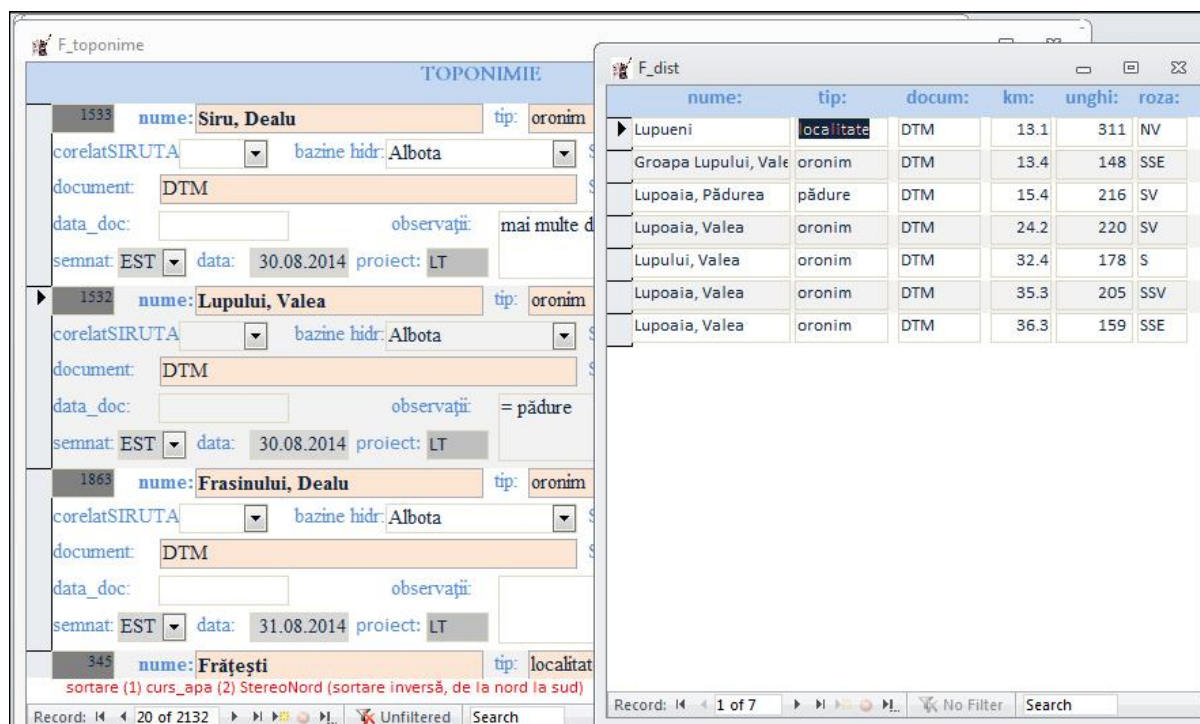

Figure 4.2. The function of analogies seeking.

---

[2] It was not a theoretical idea at all, but it occurred at the request of Bogdan Şandric, from the National Heritage Institute, who wished to expand the territorial limits proposed by this project.

The database includes a series of facilities, like the search of a name, starting from a certain location (dependent on the positioning of the cursor on the toponymy form), indicating all the similar names, the direction and distance at which these analogies are (fig. 4.2). Actually this is an application which simulates a GIS capability, applicable here also to those who does not master the use of a GIS software.

Attached to the database there is also a dictionary of words connected to toponymy, in which there are definitions of some words that are no longer used in Romanian language (fig. 4.3). In the large, at least half of the used word stock of toponyms derive from provincial words or decayed words, which are no longer known by the average speaker of Romanian language. Understanding the real signification of less transparent words is part of the effort of reconstructing the circumambience, for the strip of ground where the Romans have established their frontier, in the first half of the 3rd century.

Figure 4.3. Dictionary of rare words.

Complementary to the „Terminology" form (or Dictionary), there is an additional form, destined to strictly etymological comments.

Another helpful function of the database is the bibliographic database, which can be used for any other necessity implied by the documentation for the current research project.

Each of the database components has a „help" type file (so far Romanian only), which can be accessed directly from the application's main switch panel.

The toponymic database will be soon available (during the winter of 2014-2015) to the large public, with reading rights; the rights to write or modify are reserved to the members of the project, including the registered volunteers. We wish not only to give back to the society that subsidizes us, a gain of our work, but we also wish to use the linguistic experience of our readers who – why not? – might be able to explain words that we haven't found in dictionaries. On the site of the project there will be email addresses where our readers may send their feed-back to the issues we have raised. From them we will also choose the volunteers who will be granted the right to write in the toponymic database.

It may be useful to mention that this will be the first toponymic database available to the large Romanian public. Even if it will be limited to approx. 6000 km^2 (approximately the dimensions of a county), it will present the benefit of showing the historical dynamics of names. The database will be completed with all the toponyms discovered in military maps from various periods of time, from the

most recent (the map from the 80's, last century), to the oldest (like the Specht map, from 1791)[3]. We have no doubt that many specialists – who are not members of this project – will find in it a source of inspiration for their professional activity, be they historians or linguists.

The statistical analysis of the sets of toponyms will be made on several extension sections: the parishes' current territory; hydrographic basins; current counties (about one third each). The purpose of the analysis is to observe local/regional particularities, in the hope that we will be able to mark out the most characteristic features of the surroundings.
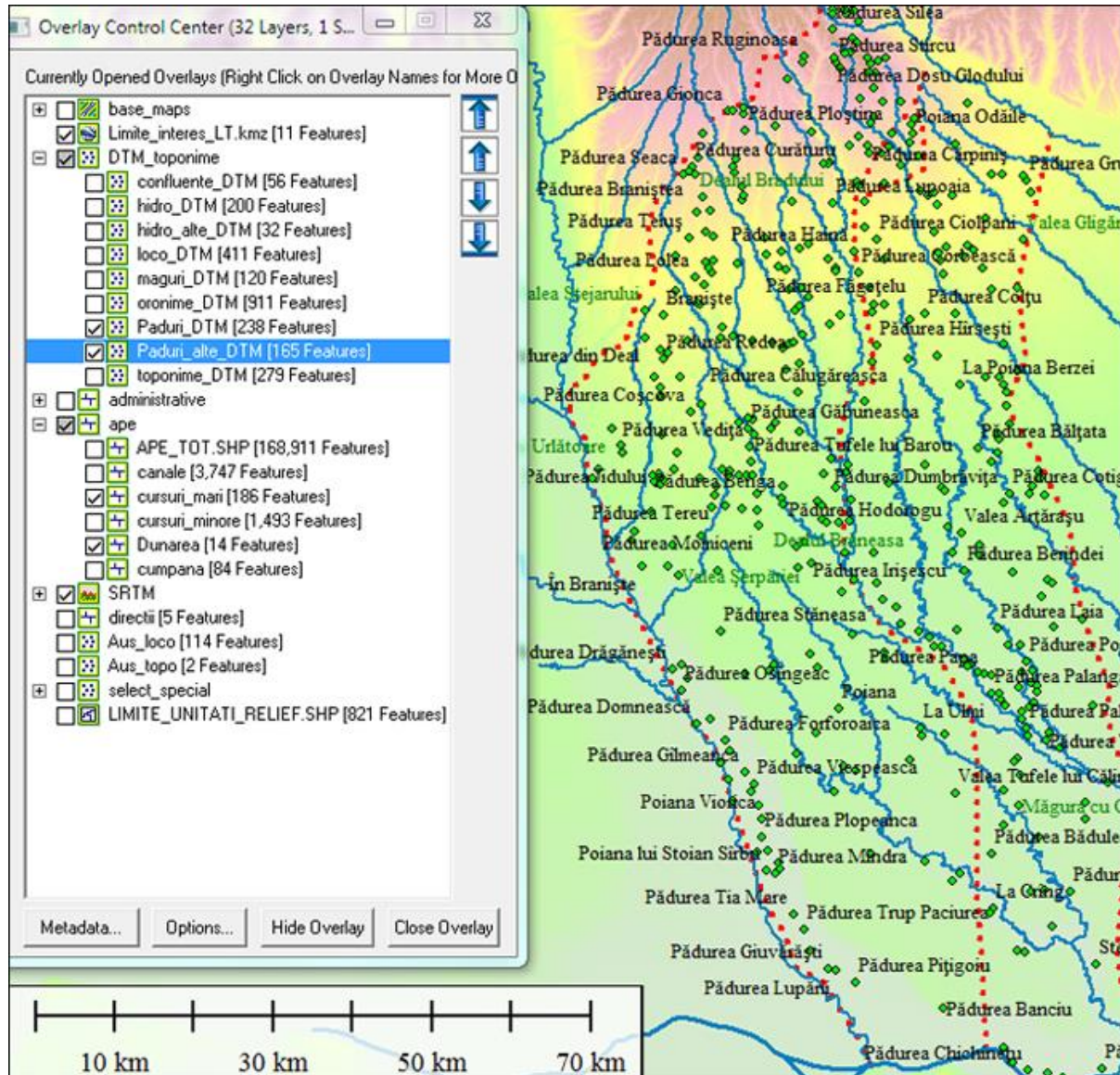

Fig. 4.4. Illustration of an operating window in Global Mapper.

The dotted red lines represent the territorial operational limits of this project.

The database we briefly described above is the replica of a GIS file, which contains the same data, in the specific format (fig. 4.4). The reason why a database was also necessary is that not all the members of the project can handle a GIS, nor all of the future beneficiaries of our work. As we have

---

[3] In intermediary positions are Planurile Director de Tragere (made between 1916 and 1966), a Treia Ridicare Austriacă (collected after 1868, but published in 1910), and the so-called Szathmári map, with data gathered in 1855-58, published in 1864.

shown, the database allows the development of a linguistic interface, by means of which we will try to riddle out names which are not always transparent, an aspect which is harder to solve in a GIS environment.

The organisation of a GIS file is the usual one, working with raster files (all the historical maps from which data is collected, but also SRTM[4] files, which allow three dimensional representations of the territory), or with vector type objects, be they imported (for instance public limits), or locally produced (layers of data for toponymy).

The data produced in the GIS file are exportable in data libraries which may easily become public assets, by means of some popular format, like Google Earth.

---

[4] Acronym for Shuttle Radar Topographical Mission, SUA, 2000.