

## **Thematic issue**

# **Historical Databases (1)**

### **Editor's Note**

The first historical population databases, explicitly dedicated to the digitalization of information extracted from parish registers and censuses prior to the First World War, were created in the 1970s in northern Europe (Brändström 1996). Over the following decades, with the growing entanglements between historical research and computer technology, these tools became increasingly popular in North America and western and northern Europe, culminating in the early 2000s with the outset of the North Atlantic Population Project (NAPP) (Roberts et alii 2003). NAPP continues to this day, aiming to provide data integration and support for wide-scale analyses based on the population databases of the North Atlantic regions ([https://www.nappdata.org/napp/revison\\_history.shtml](https://www.nappdata.org/napp/revison_history.shtml), accessed 15.06.2020), its contributors have also initiated the development of an intermediate data structuring system designed to enable the integration and harmonization of data from other similar historical sources (IDS) (Alter, Mandemakers and Gutmann 2009; Alter and Mandemakers 2014).

A second wave of historical population databases development has covered, during the last decade, southern and central Europe and East and Southeast Asia (<https://ehps-net.eu/databases>, accessed 15.05.2020) rooted in the methodological foundations laid by the aforementioned predecessors, and encouraged by the spectacular scientific results achieved by employing digital tools. The accelerated technological progress, which lowered the costs of such enterprises, also played an important part. Although many of the new databases only cover limited areas or narrow chronological periods, they remain fundamental instruments in historical demographical research and in the study of early modern, modern and contemporary populations.

Within this international framework began the construction of the “Historical Population Database of Transylvania” (HPDT) in 2014, through a project funded by European Economic Area and Norway Grants, involving cooperation between the Centre for Population Studies at Babeş-Bolyai University Cluj-Napoca and the Norwegian Historical Data Centre at the University of Tromsø ([http://hpdt.ro:4080/docs/documents/docs/5/original\\_Utilizarea\\_bazei\\_de\\_date.pdf](http://hpdt.ro:4080/docs/documents/docs/5/original_Utilizarea_bazei_de_date.pdf), accessed at 15.06.2002). The initiative befitted the long tradition of historical

demographical research at Babeş-Bolyai University, which reaches back to the beginnings of the institutionalization of historical demography as a science, in the 1960s and 1970s (Bolovan and Bolovan 2003). These early steps were followed by an increased interest and activity in the 1990s and 2000s, impelled by the establishment of the Center for Population Studies (2003), which furthered contacts with the scientific community by means of publications, scientific events and membership in international forums.

HPDT is a relational database covering micro-areas of the territory of historical Transylvania (the former Habsburg province of Transylvania before 1867) for the period 1850–1914. The hosted data comes mostly from the parish registers preserved at the county services of the National Archives of Romania or in local parishes. HPDT has been operating since 2015, and its first stable version was launched in 2017, following the completion of the original project. Its genealogical interface (<http://hpdt.ro:4080/>, accessed 15.06.2020) was also made public at the same time. During the following years, its construction and technical adaptations to the challenges raised by historical documents have continued, as well as the input and processing (verification, cleaning, standardization, linkage) of information from primary sources, including the publication of the documentation and of the first scientific results, which are detailed in the study herein.

The present issue of the journal “*Studia Universitatis Babeş-Bolyai Digitalia*” features a special, more discreet piece of the HPDT’s documentation, namely the English version of the manual for transcribing and inputting information from historical parish registers into the database entry-forms. Publishing the data-entry protocol is not a common option in the scientific literature, and even on databases’ websites such documents are less frequent. Indeed, this manual is part of the internal documentation, employed mainly by the data-entry operators and by the database core-team when conducting verification and normalization procedures. However, it remains an important item, both in regard to the digital instrument for which it was created, and to other emerging initiatives in search of methodological benchmarks.

During the early stages of HPDT’s construction, the identification of accessible models for creating this manual proved challenging. Moreover, the provincial specificities create a quite different pattern from the ones implied by the standards encountered in northern and western Europe, including an overall less satisfactory level of sources’ preservation and informative content, a wider confessional and linguistic diversity, lower levels of professionalization of the church personnel for some of the denominations – mainly the Orthodox, sometimes the Greek Catholic –, and some of the lowest levels of general literacy in the Habsburg Monarchy. All of the above required the accommodation of the database structure and data-entry procedures to the forms and contents of the sources, resulting in a mandatory, low to moderate, level of normalization during the latter’s transcription process.

The manual details the balance between perfectly replicating the information from the sources (e.g. anthroponomy and toponomastics, nobility and occupational titles, cause of death), and the employment of normalization procedures (e.g.

timestamps, age, sex, denomination), together with the structural adaptations meant to accommodate information from extremely diverse and chronologically dispersed sources into a small number of standard digital forms. The historical diversity of Transylvania being characteristic, to a certain extent, of the whole central and south-eastern Europe, and given that, for the time being, HPDT remains a unique tool in the region in terms of technical achievement and complexity, the solutions used in its development can prove useful for similar digital instruments covering neighboring areas, or for other historical topics. On the same note, it is only normal that the auxiliary of such a large digital tool, with a potential to serve as a benchmark, should be made available to the scholarly community for exploration and review.

On these grounds, the project team deemed it necessary to publish a translated version of the data-entry manual in an open-source journal, and the editors of "Studia Universitatis Babeş-Bolyai Digitalia" welcome this initiative.

Vlad POPOVICI

### Works cited

- Alters, George and Kees Mandemakers. "The Intermediate Data Structure (IDS) for Longitudinal Historical Microdata, version 4." *Historical Life Course Studies*, vol. 1, 2014, no. 1, pp. 1-26, [https://ehps-net.eu/sites/default/files/hislives\\_vol\\_1\\_23-05-14.pdf](https://ehps-net.eu/sites/default/files/hislives_vol_1_23-05-14.pdf), 15.06.2020.
- Alters, George, Kees Mandemakers, Myron P. Gutmann. "Defining and Distributing Longitudinal Historical Data in a General Way Through an Intermediate Structure." *Historical Social Research / Historische Sozialforschung*, vol. 34, no. 3, 2009, pp. 78-114.
- Bolovan, Sorina, Ioan Bolovan. "Ştefan Pascu şi demografia istorică în România." *Anuarul Institutului de Istorie « George Bariţ » din Cluj-Napoca. Series Historica*, vol. 42, 2003, pp. 145-153.
- Brändström, Anders. "The demographic database at Umeå university in Sweden." *Annales de Démographie Historique*, vol. 34, no. 2, 1998, pp. 197-214, [https://www.persee.fr/doc/adh\\_0066-2062\\_1999\\_num\\_1998\\_2\\_1945](https://www.persee.fr/doc/adh_0066-2062_1999_num_1998_2_1945), at 15.06.2020.
- Roberts, Evan, Steven Ruggles, Lisa Y. Dillon, Ólöf Garðarsdóttir, Jan Oldervoll, Gunnar Thorvaldsen, Matthew Woollard. "The North Atlantic Population Project. An Overview." *Historical Methods. A Journal of Quantitative and Interdisciplinary History*, vol. 36. No. 2, 2003, pp. 80-88, <http://users.hist.umn.edu/~ruggles/napp.pdf>, at 15.06.2020.

