

Lyudmila Mazur, Oleg Gorbachev

## **Historical and Demographic Database: Russian Experience and Prospects**

In the paper, the Russian experience of creation of historical and demographic databases is generalized. The principles of their formation and the source base analyzed. In the context of the Russian and global experience of electronic resources development the concept and intermediate results of implementation of the "Ural Historical Electronic Archive" project in Ural Federal University are submitted.

Information base of historical research is rapidly changing due to the spread of technology. The most notable changes have occurred in historical demography. The implementation of international projects to create Big Data Base (IPUMS-USA, NAPP, Mosaic, Viennese Database on European Family History Church books, Demographic Database Umeå et al.) destroys national framework research practices and opens possibilities for development of cross-national subjects on family history, processes of fertility, mortality, nuptiality, and migration. As demographic factor is basic for political, economic, and ethnic processes, historical and demographic databases will be relevant for other areas of historical scholarship.

Ability to work with Big Data Base in remote access mode by means of different search methods, and data processing using statistical package allows the researchers to model their own resources on this basis, taking into account the goals and objectives of their research. Consequently, it is possible to speak not simply about the new informational environment of historical science, but also on new algorithms and technology-oriented research work with mass primary data.

In Russian historiography, the first experience of creating databases built on the mass historical data took place at the beginning of the 1980s. With the formation in 1992 the "History and Computer" Association, activities of Russian historians to develop the database has increased markedly.

Most actively development of this problem was conducted by Laboratory of Historical Informatics at Moscow State University. On its site, two full text databases executed by T. Izmestyeva and T. Valetov are now presented.

The most interesting results have been achieved in the field of historical demography.

In 1990-2000s, several large projects were realized. In 1997-2005 in the Saint-Petersburg State University, the group under the guidance of prof. S. G. Kashchenko created a DB on metric books of Olonets province of 18 - early 20 century.<sup>1</sup>

Somewhat later historians from the Tambov State Technical University led by prof. V. V. Kanishchev addressed to the same source.<sup>2</sup> Revision lists and inventory homestead became an object for modeling databases on several districts of Penza province in 18–19 century.<sup>3</sup> Another center of creation of historical and demographic databases became Altai State University (Barnaul). Here preserved primary materials of the first Russian Empire census (1897) were found, and served as the foundation of the database, supplemented with the materials of parish records.

Thus, within a historical demography the special direction presented by several regional centers which can be designated as Digital Historical Demography was created. Formation of electronic resources based on primary historical sources became its feature. All projects had local character, i.e. covered certain settlements, the districts of Russia of 18–19 centuries.

Databases were developed as specific research projects and did not become the common property, for an exception of an informational system "Historical Occupational studies" created under the guidance of prof. V. N. Vladimirov. It included the "Population of Barnaul in XIX —

early XX Century” database built on births registers, outputs of 1897 census, and other mass sources.

Thus, creation of the national project is not scheduled yet though prerequisites for this purpose already developed, as well as need for it. For example, in 2011 the research project of Higher School of Economics "Development of the Database of demographic indicators on regions of Russia and on the countries of the world" which pursued the aims of association of the major demographic indicators across Russia and other countries and providing access to the saved-up information in one database in a uniform, convenient format via the flexible web interface was started. The objectives of the study was describing the sources of demographic information, analyzing available databases, identifying the characteristics of their interfaces, accuracy, format and other criteria. One of research problems was the description of sources of demographic information, the analysis of the available DBs, detection of features of their interface, accuracy, formats and other criteria.<sup>4</sup> It is one of the first projects focused not on the internal purposes and tasks, but on the external research environment and its informational support.

Resulting in the implementation of these projects, the experience proved to be very useful to clarify the concepts of the resource environment for creating databases. All researchers note incompleteness of preserved complex of primary documents, their fragmentation, complexity of processing. The majority of primary documents scattered in local archives in the Russian regions; preserved complexes are not always comparable on registration technologies and record structure. Nevertheless, positive experience of formalization of historical sources data and development of algorithms of their further processing was saved up.

Given the existing foreign and domestic experience of creating public historical database, in the International Demographic Unit (IDUn) of the Ural Federal University in 2014 development of the regional network resource "Ural Historical Electronic Archive" began. Its main objective is the formation of diversified historical electronic archive and full-text databases with digital library and archival documents, maps, photos, video documents (e-Archives) on regional history based on the mass sources preserved in the local archives.

In the first phase (2014-2015) several databases were created: 1) based on some parish registers of Yekaterinburg diocese for 1800-1918; 2) based on the Communist Party census 1922-1924; 3) DB "1959 census in Sverdlovsk region"; 4) DB "Population of urban and rural settlements of the Middle Urals in 19–21 century"; DB "Circumpolar Census 1926-27". In the next phase (2016-2017) it is planned to expand the available arrays through digitization and decryptization of census lists of Perm province, as well as the household's registers.

It is necessary to allocate those basic principles on which developers of the project rely:

- 1) focus on the primary mass historical sources (population censuses, parish books, and other variants of population registers) that contain personal information
- 2) continuous coverage of all extent of documents and records
- 3) use of the data format providing store, import, integration of multi-temporal resources and their collective use
- 4) the openness and accessibility of resources created by the scientific community.

Given the incomplete and fragmentary preservation of many documents in Russian archives, including primary census forms, there is a need to formalize and digitize of all extant multi-temporal set of documents on the history of the Urals of 18–19, and early 20 centuries.

As practice shows, overcoming the fragmentation of research is possible. For several years already, the historians and demographers from a number of countries are successfully

implementing the European Historical Population Samples Network (EHPS-Net) project, which aims to develop a common format for Intermediate Data Structure (IDS) for more than 20 population databases. The inclusion of the Ural Historical Electronic Archive to EHPS-Net will allow correlating Russian demographic processes with European and international, determination their specificity and prospects. In the case of integration with IDS, main hopes of Russian researchers will be associated with the study of vital statistics and the evolution of family forms, as well as the specifics of ethno-religious communities.

### **Funding**

This work was supported by the Russian Foundation for Basic Research [grant number 15-06-05611].

### **References**

---

1. Izuchenie istoricheskoy demografii Rossiyskoy imperii na kafedre istochnikovedeniya istorii Rossii Sankt-Peterburgskogo gosudarstvennogo universiteta [The study of historical demography of the Russian Empire at the Department of Source History of Russia, St. Petersburg State University] (2013). *Informatsionnyy byulleten' Assotsiatsii «Istoriya i komp'yuter»*, 41, 137 – 141.

2. Kanishchev, V.V., and Konchakov, R.B. (1999) Puti sozdaniya i obrabotki istochniko-orientovannoy bazy dannykh na materialakh metriceskikh knig (prikladnoye s. Baykalovka Tambovskoy gubernii, 1810 – 1918 gg.) [The creation and processing of source-oriented database on metric books materials (parish of Baykalovka, Tambov province, 1810 - 1918)]. *Krug idey: istoricheskaya informatika na poroge XXI v. Moscow-Tcheboksary*, pp. 184-189.

3. Arsent'yev, N. M., Pershin, S. V., and Shcherbakov V.V. (1997). *Novye gorizonty v issledovanii revizskikh skazok: iz opyta sozdaniya banka dannykh «Region»* [The new horizons in research of the census lists: from experience of creation of the "Region" databank]. *Krug idey: traditsii i tendentsii istoricheskoy informatiki. Moscow*, pp. 139 – 145.

4. Vinnik, M. V. *Kakoy dolzhna byt' ideal'naya demograficheskaya Baza dannykh?* [What has to be an ideal demographic Database?]. *Demoskop. Derived from:* <http://demoscope.ru/weekly/2012/0533/analit06.php>