
GIS in gathering information from historical sources

Michał Gochna*¹ and Grzegorz Myrda*¹

¹The Tadeusz Manteuffel Institute of History, Polish Academy of Sciences – Poland

Abstract

Amongst many types of historical sources used in research of history of Poland from late medieval times to the the end of 18th century some have fortunately preserved in a large amount. Those are mostly court books and fiscal sources. Although they bring a significant information regarding many political, social and economic aspects of history, their massive-ness brings major difficulties in obtaining necessary information. Many attempts have been made to simplify the access to content of those sources. In case of fiscal documents, full-text editions of tax registers and various kinds of bills and receipts has been prepared. Elaboration of court sources, which are mostly books of court records, vary from full-text editions in the case of the oldest books, to the summaries of content in the case of later books. Other way of making historical massive sources accessible, which is being more and more popular, is to prepare scans of manuscripts and put them on the Internet with an open access. An example of the latter can be found on the website "Search in Archives", where the Polish State Archives successively publish new scans of documents. In this approach, however, one can work only with a simple scan, equipped with basic metadata (storage place, signature, title of the unit). Acquiring data from these materials requires reading the source content in the same way as in the case of working with originals in the archive. In other words, two main approaches of publishing massive historical sources could be lined out. On the one hand it is a traditional, full-text edition or preparations of summarizes, both, in most cases, requiring a significant amount of time of research work; on the other hand, a simple on-line publication of the scans of manuscripts. While the first approach is valuable in publishing sources of limited size, it does not solve the problem of facilitating the use of mass sources. The publication of scans, in turn, changes only the form of access to the source and the medium on which it is presented. The problem of developing a larger portion of material in a reasonable time still last.

A proposition of solution of this issue has been developed during the preparation of the projects regarding the historical geography of Poland that were, and still are, conducted in the Tadeusz Manteuffel Institute of History of Polish Academy Of Sciences. Two types of digital editions of historical sources were prepared: a digital edition of tax registers from the Greater Poland from the 2nd half of the 16th century, and a digital edition of Greater Poland court books from 16th century. Both of them were prepared with the use of GIS (Geographical Information Systems) tools and/or GIS technology, as those can be seen as universal platform for collecting, processing, analyzing and sharing data from different sources and of different kind, including geographical, but also non-geographical information contained in historical manuscripts.

The use of standards existing in GIS world, resulting from natural desire to reduce the costs of time-consuming data gathering used by many different groups of recipients, contributes to increased interoperability when it comes to utilising collected information.

*Speaker

Obtained data stored in database is software agnostic, however compliant with hundreds of software products that are already implementing or have been certified with Open Geospatial Consortium standards, so can be used in wide range of applications for further processing.

According to above rules some software during different projects was developed to facilitate the process of entering specific data, like digital editing of historical manuscripts application called INDXR or OntoForm aimed at helping in building database driven by ontology.

Foundation for all of above solutions is common database, accessible through the Internet, for everyone interested in using it. The data collected can be expanded, updated, verified using dedicated web tools associated with database or any GIS application to which the user is accustomed to. This makes it much easier to achieve cross-project synergies using the same collected data, or enriching it with additional layers of information.

Work were carried out by historians, geographers, ontologists, computer scientists, who were trying to find theoretical foundations and methods of collecting source information in a way that allows integration of data descending from different sources, different periods of time, different areas, and having different content.

Keywords: massive sources, digital source edition, GIS, HGIS, history, historical geography, Big Data