Integrating Historical Census Data in the Semantic Web

Albert Meroño-Peñuela[1,2] and Ashkan Ashkpour[3,4]

[1] Department of Computer Science, VU University Amsterdam, NL
[2] Data Archiving and Networked Services, KNAW, NL
albert.merono@vu.nl
[3] Erasmus University Rotterdam, NL
[4] International Institute of Social History, NL
ashkan.ashkpour@iisg.nl

Abstract. Historical censuses are one the most consulted, reliable and large scale statistical data sources available, describing the demographic, social and economic history of a nation. To answer their research questions, researchers often need to query long time series of census data. However, such longitudinal queries are typically hampered by the scarce integration of the historical censuses, demanding manual and knowledge intensive harmonization and restructuring in order to obtain meaningful comparisons over time. The challenges are even harder if provenance microdata is lost. In this paper we describe the methdology followed in CEDAR[5], a project of the Computational Humanities Programme[6], to provide solutions to these data-issues in the Dutch historical censuses (1795-1971). Our proposal builds on top of Linked Data and the Resource Description Framework (RDF) technology, allowing us to transform the original census tables into a graph of Linked Census Data. With such a graph, every census data-item can be interlinked on the Web with other hubs of historical socioeconomic and demographic information. By following the Linked Data principles, our aim is two-fold. On the one hand, we show how the integration of our own historical census data is improved by linking them to the network of Linked Historical Datasets on the Web. On the other hand, we envisage new historical classifications (like demographical structures, housing types, occupational classes and statuses, or religious denominations) coming out of our harmonization process, which are not published yet on the Web on a standard manner and could improve the interoperability of other datasets.

---

[5]See http://www.cedar-project.nl/
[6]See http:///www.ehumanities.nl/