A visual historical exploration tool for lexical resources: The case of Austrian language

Alejandro Benito, Antonio Losada, Roberto Theron,
Eveline Wandl-Vogt, Amelie Dorn, Melanie Seltmann

Digital humanities usually follow classical approaches while researching for new exploration methods. The state of the art of the field covers works where the text exploration is the main means to discover knowledge, leaving other tools such as those based on data visualization aside. In this paper we introduce into the project "exploreAT! exploring Austria´s culture through the language glass" (2015-2019) by focusing the visual analysis workpackage.

We explain the creation of a visualization system based on the Database of Bavarian Dialects in Austria (DBÖ, dbo@ema). In it around 2,200,000 records of lemmas related to the Austrian corpus can be found, several of them part of a unique data collection (1920-1950), which were registered during the past 500 years (of which 232,000 contains geographical and temporal information). This dataset is stored in XML files where each record presents multiple fields, those having diverging formats. Our proposed prototype is a multidimensional, visual analysis tool that allows exploration of the data. Despite supporting several dimensions, the analysis workflow is guided by the spatial dimension and thus the user interface is greatly based on the use of maps, projections and other visualization artifacts (communities graphs, timelines, colocation graphs) built on top of them. We discuss the key role visualization plays releasing the power of mathematical and computational methods (such as network analysis, natural language processing and geographic information systems) that is likely to be a valuable contribution in the field of visual DH in general, and lexicography and cultural (heritage) studies in particular.