Abstract

The aim of the thesis is to present an overview of topics that were the subject of research of the Czech Academy from 1890 to 1910 using computational analysis of its publications. The work complements and extends the existing state of knowledge in this subject area, which has so far been limited to analyses of specific publications, fields of study or scientific classes of the Academy, but did not allow a view of the areas of interest of the Academy as a whole. The introductory part of the thesis presents the context of the Academy's founding, summarizes the current state of understanding of the topics that Academy scholars have addressed and describes how the Academy's digitized publications can be used for analysis using computational methods within the digital humanities. In the research part, digitized issues of the scholarly journal *Rozpravy*, which represented the core publication platform of the Czech Academy, were obtained from the Digital Library of the Czech Academy of Sciences. The obtained data were processed using freely available tools and used as input for topic modelling using the LDA (Latent Dirichlet Allocation) method. The result of the thesis is a comprehensive overview of 35 specific topics that the Czech Academy addressed in the first twenty years of its existence. It describes the number of publications under each topic and their changes during the studied period. Also, the semantic similarity of each topic and its relation to the present-day categories of science classification is presented. The thesis fulfills its aim and provides a new perspective on the publication activity of the Czech Academy. It briefly outlines possible further research directions consisting in the expansion of the corpus of publications studied or a more detailed evaluation of the topic model. The thesis also shows that documents available in digital libraries can be used for research in the digital humanities, although this is not their primary intended use.

Keywords

Czech Academy of Emperor Franz Joseph for Science, Literature and Arts; topic modelling; LDA; digital humanities