The development of the information science research in Croatia: Analysis of scientific paper title words

Mirjana PEJIĆ BACH[a,1] Jadranka STOJANOVSKI[b,c] Jasmina PIVAR[a] Jovana ZOROJA[a]

[a] *University of Zagreb, Faculty of Economics & Business*
[b] *Ruđer Bošković Institute;* [c]*University of Zadar, Department of information science*

Abstract. Titles of scientific papers convey crucial topics of the articles to readers, reviewers or editors, and give a very compact summary of the publication. They can be analyzed to gain insights into the current and past interests and trends of scientific work of scientists in different fields. Our goal is to identify common topics from the discipline of library and information science (LIS) in order to gain insight into recent research in Croatia in this area. Two primary research questions are: (a) is there the difference between LIS research topics of Croatian researchers and common LIS topics globally, and (b) can the word clusters gained from the topics of the research in LIS area be compared with the popular LIS taxonomies?
In our study, we analyze the titles of the scientific papers published by Croatian scientists covering the field of information sciences. Croatian Scientific Bibliography (CROSBI) provided the data for the purpose of our study. CROSBI was launched in 1997 and has evolved into a comprehensive bibliography. It collects all the literature produced by Croatian scientists: journal articles, conference papers, books, book chapters, reports, manuscripts, etc. Authors provide the data into the database by themselves, and the bibliography is monitored by librarians. The structure of the database and rich set of metadata for every publication type makes CROSBI bibliography suitable for many purposes, including different types of content analysis. For the purpose of this study we examine the titles of scientific papers published during the period from 1997 to April 2015.

After the introduction, we are going to describe the methodology of our study, including data collection and data analysis. First, we are going to describe the process of journals and papers selection. Then, we are going to identify the most frequently occurring words and phrases. Stop words and non-specific words will be treated by using the usual practice of excluding them from the further analysis. The titles of scientific articles will be analyzed by using few methods of word analysis to fulfill the goals of our study. First, we are going to use relative word frequency with the goal to describe the distribution of term occurrences across the titles of the scientific papers. Second, we are going to use co-word analysis and hierarchical clustering to determine clusters of closely related terms appearing in the titles of the scientific papers. Our results will be compared with the results from the random sample of the titles from Library, Information Science & Technology Abstracts (LISTA) database as well as with different classification schemes of the LIS field. Furthermore, we are going to describe and discuss the results of our study in details. At the end, the brief conclusion the main findings of our study are going to be provided.

---

[1] Corresponding Author.