# Space of articles' thematic proximity modeling using fuzzy approach

Aleksandra MRELA[1], Oleksandr SOKOLOV[2], Veslava OSIŃSKA[2] and Mateusz KANIECKI[2]

[1]Kazimierz Wielki University in Bydgoszcz, Poland
[2]Nicolaus Copernicus University in Torun, Poland

Correspondence should be addressed to: Aleksandra MRELA; a.mrela@ukw.edu.pl

*Presented at the 42nd IBIMA International Conference, 22-23 November 2023, Seville, Spain

## Abstract

The thematic similarity of scientific articles can be examined using bibliographic metadata such as titles, keywords, descriptions, and distribution of citations. Researchers use traditional co-citation analysis, bibliographic linking, or text analysis to process and further organize publications. Scientometric works have reported that they put more and more effort into explicit combinations of the different approaches to obtain and verify the achieved results. These studies show that the data quality strongly determines the chosen method. If we lack a valuable, representative data set, the researchers' attention leads toward modeling an appropriate metadata space. Hence, academics can construct a suitable semantic space if they need to organize a set of publications.
We propose a method using two factors for searching for the most appropriate articles from a set of publications dealing with similar issues. One measure is based on commonly used keywords, while the other takes into account articles that the selected person cites or is cited by. We use these factors to create a thematic proximity space, showing some distance from the first studied publication. Furthermore, we define a Takagi-Sugeno fuzzy inference system for publication classification.
The proposed model of the thematic proximity space based on keywords and article citations can be the basis for an information search interface dedicated to researchers. They can use it to search and filter relevant articles. Fuzzy logic rules used to model this space are a new concept in this type of research.

**Keywords:** thematic proximity space, keywords and citation analysis, fuzzy approach