

Logical structure analysis of scientific publications in mathematics

Solovyev V., Zhiltsov N.

Kazan Federal University, 420008, Kremlevskaya 18, Kazan, Russia

Abstract

Even though the Linking Open Data cloud is constantly growing, there is a serious lack of published data sets related to the domain of academic mathematics. At the same time, since most scholarly publications in mathematics are well-structured and conventional, it's promising to get their helpful detailed representation. The paper describes an approach to extracting and analyzing the structure of mathematical papers. We present the Mocassin ontology that is used by analysis algorithms and can be considered as an ontology of the structure of scholarly publications in mathematics. The proposed semantic model has been evaluated on a set of real mathematical papers and preliminary evaluation results are encouraging. Also we discuss potential applications of the model to specific information retrieval tasks including semantic search. © 2011 ACM.

<http://dx.doi.org/10.1145/1988688.1988713>

Keywords

document structure analysis, linked data, ontology