

Ontological-semantic text analysis and the question answering system using data from ontology

Kuznetsov V. A.*, Mochalov V. A.** , Mochalova A. V.**

* *Petrozavodsk State University, Lenina str. 33, 185910 Petrozavodsk, Russia*

** *Institute of Cosmophysical Research and Radio Wave Propagation FEB RAS, Mirnaia str. 7, 684034 Paratunka, Kamchatka region, Russia.*

kuznetcv@mail.ru, sensorlife@mail.ru, stark345@gmail.com

Abstract—Today, with an avalanching increase in information, the task of developing the systems that allow user to quickly search desired information in large text volumes is becoming more and more urgent. An example of such system is the question answering one. In the work we describe an architecture of such system which work is based on utilizing data from an ontology. We propose an algorithm for automatic update of the ontology basing on use of an expert system and ontological rules for logical inference. We also describe an ontology with the structure based on object-oriented model and describe the functions that are used to update the ontology and extract data from it. We describe the way to update the ontology and to modify the stored data using the rules stored in the ontology. For writing ontological-semantic rules we use the Drools expert system that utilizes the PHREK algorithm for fast pattern matching. We analyze the issues of using Apache Spark system for distributed implementation of the algorithm.

Keyword— Question Answering System, ontology, expert systems, semantic analysis.



Vladimir Kuznetsov is a scientist, mathematician, Professor, Doctor of technical sciences, Honorary worker of higher professional education of the Russian Federation, Honored worker of science of the Republic of Karelia. He is the author of over 200 scientific papers in the field of mathematical modeling and solution of applied optimization problems, as well as academic literature on the programming. He is one of the coaches of the teams who were winners of the world championship on programming (in 2007, 2008, 2010). His research interests include optimization techniques, Olympiad programming.



Vladimir Mochalov was born in Lyubertsy, Russia in 1985. He received the Ph.D. degree in electronic engineering from Moscow Technical University of Communications and Informatics. His research interests include networks structure synthesis, artificial intelligence, bio-inspired algorithms, query answering systems and Big Data.



Anastasia Mochalova was born in Petrozavodsk, Russia, in 1987. She received the bachelor's degree at Petrozavodsk State University, the master's degree in St. Petersburg State University of Aerospace Instrumentation. She is an external PhD student in technical sciences at Petrozavodsk State University. Her research interests include automated processing of natural language texts, development of question-answering systems, automation of ontologies creation, development of the semantic analyzer.