

Synthetic Method of Contextual Knowledge Extraction

Authors : Olga Kononova, Sergey Lyapin

Abstract : Global information society requirements are transparency and reliability of data, as well as ability to manage information resources independently; particularly to search, to analyze, to evaluate information, thereby obtaining new expertise. Moreover, it is satisfying the society information needs that increases the efficiency of the enterprise management and public administration. The study of structurally organized thematic and semantic contexts of different types, automatically extracted from unstructured data, is one of the important tasks for the application of information technologies in education, science, culture, governance and business. The objectives of this study are the contextual knowledge typologization, selection or creation of effective tools for extracting and analyzing contextual knowledge. Explication of various kinds and forms of the contextual knowledge involves the development and use full-text search information systems. For the implementation purposes, the authors use an e-library 'Humanitarianiana' services such as the contextual search, different types of queries (paragraph-oriented query, frequency-ranked query), automatic extraction of knowledge from the scientific texts. The multifunctional e-library «Humanitarianiana» is realized in the Internet-architecture in WWS-configuration (Web-browser / Web-server / SQL-server). Advantage of use 'Humanitarianiana' is in the possibility of combining the resources of several organizations. Scholars and research groups may work in a local network mode and in distributed IT environments with ability to appeal to resources of any participating organizations servers. Paper discusses some specific cases of the contextual knowledge explication with the use of the e-library services and focuses on possibilities of new types of the contextual knowledge. Experimental research base are science texts about 'e-government' and 'computer games'. An analysis of the subject-themed texts trends allowed to propose the content analysis methodology, that combines a full-text search with automatic construction of 'terminogramma' and expert analysis of the selected contexts. 'Terminogramma' is made out as a table that contains a column with a frequency-ranked list of words (nouns), as well as columns with an indication of the absolute frequency (number) and the relative frequency of occurrence of the word (in %% ppm). The analysis of 'e-government' materials showed, that the state takes a dominant position in the processes of the electronic interaction between the authorities and society in modern Russia. The media credited the main role in these processes to the government, which provided public services through specialized portals. Factor analysis revealed two factors statistically describing the used terms: human interaction (the user) and the state (government, processes organizer); interaction management (public officer, processes performer) and technology (infrastructure). Isolation of these factors will lead to changes in the model of electronic interaction between government and society. In this study, the dominant social problems and the prevalence of different categories of subjects of computer gaming in science papers from 2005 to 2015 were identified. Therefore, there is an evident identification of several types of contextual knowledge: micro context; macro context; dynamic context; thematic collection of queries (interactive contextual knowledge expanding a composition of e-library information resources); multimodal context (functional integration of iconographic and full-text resources through hybrid quasi-semantic algorithm of search). Further studies can be pursued both in terms of expanding the resource base on which they are held, and in terms of the development of appropriate tools.

Keywords : contextual knowledge, contextual search, e-library services, frequency-ranked query, paragraph-oriented query, technologies of the contextual knowledge extraction

Conference Title : ICITS 2017 : International Conference on Information Technology and Society

Conference Location : Rome, Italy

Conference Dates : December 11-12, 2017