

Editorial

Advances in Semantic Information Retrieval



Recent advances in semantic technologies have resulted in methods and tools that allow creating and managing domain knowledge. They influence the ways and forms of document representation in computer memory; they define approaches to analyze documents, and techniques to mine

and retrieve knowledge. Searching for video, voice and speech materials raises new challenging problems for information retrieval systems.

We are pleased to introduce this special issue of Informatica journal that includes four revised and extended papers, presented at the 1st International Workshop on Advances in Semantic Information Retrieval (Szczecin, Poland, September 18-21, 2011).

The first paper entitled A Query Expansion Technique using the EWC Semantic Relatedness Measure by V. Klyuev and Y. Haralambous proposes the usage of a new relatedness measure in the query expansion task. One of the goals of query expansion is to reformulate the user query in order to reduce the number of non-relevant documents retrieved by search systems. EWC extends the gaining popularity explicit semantic analysis (ESA). EWC stands for ESA, plus Wordnet, plus Collocations. Whereas ESA takes into account only encyclopedic knowledge from Wikipedia, EWC considers encyclopedic, ontological, and collocational knowledge about terms. The authors use this advantage of EWC over ESA to find more precise terms to expand the user queries. They tested proposed techniques on the NTCIR data collection which is similar to that of TREC. The authors discuss the details of proposed technique, investigate the nature of improvements in the retrieval performance, and outline the directions for the future experiments.

The next paper by A. Patyk-Łońska, M. Czachor, and D. Aerts entitled Distributed Representations Based on Geometric Algebra: the Continuous Model introduces a new model of distributed representations of complex structures (sentences in natural languages) based on geometric algebra. They compare it with two other models: Holographic Reduced Representation and Binary Spatter Codes (BSC). Results of their evaluations show that the best models for storing and recognizing multiple similar statements in natural languages are the new model and BSC with recognition percentage above 90%.

The third paper entitled Experiments on Preserving Pieces of Information in a Given Order in Holographic Reduced Representations and the Continuous Geometric

Algebra Model by A. Patyk-Łońska addresses the need to develop a new scheme for encoding and decoding complex structures that is based entirely on geometric symbols. The author studies the properties of Geometric Analogues of Holographic Reduced Representations as the ability to store pieces of information in a given order by means of trajectory association. The work describes the results of three types of the experiments: finding correct item or correct place of an item in a sequence and finding the alignment of items in a sequence without the precise knowledge of trajectory vectors.

The fourth paper entitled Grammar Checking with Dependency Parsing: a Possible Extension for LanguageTool by M. Mozgovoy examines the use of dependency-based syntactic parsing in the problem of grammar checking. The author proposes a possible extension for a well-known open



source grammar checker LanguageTool. This extension will allow the users to compose new grammar rules, based on word-word dependency links. The paper demonstrates real situations, where such capabilities are helpful. The author also proposes rule syntax, similar to existing conventions of LanguageTool, and discusses implementation and testing issues.

The first year of the ASIR workshop was successful. We received 12 submissions, and 7 of them were accepted and presented on-site. We believe that the workshop will facilitate discussion of new research results in this area, and will serve as a meeting place for researchers from all over the world. Our aim is to create an atmosphere of friendship and cooperation for everyone, interested in computational linguistics and information retrieval. The workshop is now established as an event within Federated conference on computer science and information systems (FedCSIS), annually organized by the System Research Institute of the Polish Academy of Sciences and the Polish Information Processing Society, and sponsored by the IEEE.

In its turn, ASIR is supported by the University of Aizu (Japan), known as Japan's first university, solely dedicated to computer science engineering. The University of Aizu is a major center of international education and the home of several conferences, sponsored by the ACM and the IEEE.

We would wish to acknowledge selfless efforts of our committee members and FedCSIS conference organizers, who ensured high quality of publications and flawless

arrangement of the forum. We would like to specially mention professors Marcin Paprzycki, Maria Ganzha, and Halina Kwasnicka, responsible for FedCSIS. We had a great support from our international team of reviewers, consisting of: Wladyslaw Homenda, Maciej Piasecki (Warsaw University of Technology, Poland); Antoni Ligeza (AGH University of Science and Technology, Poland); Nikolay Mirenkov, Alexander Vazhenin, Ryuichi Oka (University of Aizu, Japan); Marek Reformat (University of Alberta, Canada); Qun Jin (Waseda University, Japan); Eloisa Vargiu (University of Cagliari, Italy); Tuomo Kakkonen (University of Eastern Finland); Roman Shtykh (Rakuten Inc., Japan); Slawomir Zadrozny (Systems Research Institute of the Polish Academy of Sciences, Poland); Vladimir Oleshchuk (University of Agder, Norway); Kamen Kanev (Shizuoka University, Japan); Cristian Lai (CRS4, Italy); Troels Andreasen (Roskilde University, Denmark); Anna Fensel (Vienna FTW, Austria); Evgeny Pyshkin (Saint-Petersburg State Polytechnical University, Russia); Shih-Hung Wu (Chaoyang University of Technology, Taiwan); Vladimir Dobrynin (Saint-Petersburg State University, Russia); Simone Ludwig (North Dakota State University, USA).

We also thank Professor Matjaz Gams (managing editor of Informatica), who supported the publication of this special issue.

This year, we are organizing ASIR'2012 within FedCSIS in Wroclaw, Poland. We will continue to maintain high standards of quality and organization, set by the first workshop. We welcome all the researchers, interested in semantics and information retrieval, to join our event.

ASIR Chairs.

Vitaly Klyuev, Associate Professor

Maxim Mozgovoy, Assistant Professor

The University of Aizu

Tsuruga, Ikki-machi, Aizu-Wakamatsu, Fukushima,

965-8580 JAPAN

{vkluev, mozgovoy}@u-aizu.ac.jp