
Editorial

Jinghai Rao

Nokia Research Center,
Nokia Building, 5 Dong Huan Zhonglu,
Beijing 100176, China
E-mail: jinghai.rao@nokia.com

Darijus Strasunskas

Department of Industrial Economics and Technology Management,
Norwegian University of Science and Technology,
Alfred Getz vei 3, 7491 Trondheim, Norway
E-mail: darijuss@gmail.com

Stein L. Tomassen

Department of Computer and Information Science,
Norwegian University of Science and Technology,
Sem Saelands vei 7-9, 7491 Trondheim, Norway
E-mail: steint@idi.ntnu.no

Biographical notes: Jinghai Rao is a Senior Researcher in Nokia Research Center in Beijing. Before joining Nokia, he was a Research Scientist at AOL Lab in Beijing and a Project Scientist in Mobile Commerce Lab at Carnegie Mellon University. He received PhD Degree in Computer Science from Norwegian University of Science and Technology in 2004. His research interests include ontology, semantic web, web services, context awareness and location-base services. He has published over 20 papers in international refereed journals, conferences and workshops.

Darijus Strasunskas (PhD 2006) is a Postdoc Researcher at Norwegian University of Science and Technology (NTNU), Norway. His main research interests include, but are not limited to quantitative and qualitative evaluation of novel and traditional business models, intelligent information retrieval and ontology quality aspects, as well deployment of semantic technologies for information management. His publication list includes over 40 articles in international refereed books, journals, conferences and workshops.

Stein L. Tomassen is a graduating PhD student at Norwegian University of Science and Technology (NTNU). His research interests include semantic technologies, text mining, user context, mobile applications, ontology and information quality evaluation, in general; and improvement of information retrieval quality using ontologies, in particular. His publications include more than 20 papers in international refereed books, journals, conferences and workshops.

This special issue of the *International Journal of Metadata, Semantics and Ontologies* arises as continuation of efforts to establish a forum to advance evaluation of semantic search systems. The efforts resulted in the first International Workshop on Aspects in Evaluating Holistic Quality of Ontology-based Information Retrieval (ENQOIR 2009) and this special issue.

The progress and results in the semantic search area indicate a strong impact on performance of current search systems. Existing sparse evaluations of semantic search systems report improvement compared to traditional Information Retrieval (IR) systems. However, the results lack indications whether this improvement is optimal. Yet, majority of IR evaluation

methods is mainly based on relevance of retrieved information.

Diversity of semantic search systems makes evaluation and comparison of those systems a huge challenge. Issues in the evaluation of semantic search applications are as follows. First, there is a variety of approaches having different search goals (data retrieval, ontology retrieval, service discovery, question answering, and document retrieval). Second, the search goals are pursued by different architectural and graphical user interface solutions. Third, retrieval is enhanced by processing semantics at different search phases (indexing/annotation, query processing/expansion, and, finally, ranking, organisation of retrieved information).

Such diversity advises for three objectives of evaluation:

- 1 to prove advantage in performance over existing traditional or competitive approaches
- 2 understand performance sensitivity of a system by evaluating different configurations of the system
- 3 assess usability and user experience of a system.

Standard relevance metrics are used to fulfil the first objective. However, specificity of semantic search systems requires tailored specific benchmark datasets, i.e., a set of annotated documents and relevant queries. The second objective can be pursued by development of evaluation frameworks and a set of intrinsic quality metrics, such that they would allow assessing interdependence of sub-components and deriving a ‘best-in-breed’ configuration of a system. Finally, the third objective is an ultimate goal of any system that is made for end-users. User-centric evaluation may justify costs and monetary benefits of system development. Research on measures related to user satisfaction with respect to a search system and retrieved information is a valuable effort for the community.

In order to cover this broad area a special call for papers has been issued. This special issue attracted nine submissions. After a double-blind rigorous peer review process, only three papers were selected. Two themes emerge from these papers:

- a analysis of conducted evaluation, their contemporary constraints, and a roadmap for future evaluation
- b component based evaluation – arising from willingness to understand performance sensitivity of own semantic search systems.

The first paper, by Uren et al. (2010) ‘Reflections on five years of evaluating semantic search systems’, reflects on different semantic search systems that have been developed at the KMi research group (The Open University, in the UK). The paper provides a timeline view of semantic search systems and their evaluation methods that have been adopted by the research group at KMi. The authors discuss lessons learnt in developing and evaluating search systems, and analyse a broad range of semantic search systems. Their conclusion is that scarce semantic resources have limited available evaluation methods and the consequent robustness of the performance assessments. The paper calls for efforts from the community to:

- establish benchmark datasets with regards to semantic resources and queries
- develop taxonomy of qualitative and quantitative metrics
- develop methods to examine components of the systems, e.g., logging standards.

The second paper, by Blacoe et al. (2000) ‘Evaluation of scalable multi-agent system architectures for searching the Semantic Web’, focuses on scalability issues in multi-agent based semantic search systems. Here different aspects are evaluated in a set of experiments. Namely:

- a scalability of the system has been tested in analytical evaluation of query response time with respect to query complexity, agents’ neighbourhood size, and composition
- b performance of critical components of their system is assessed, i.e., evaluating performance of semantic matching.

The third paper, by Tomassen and Strasunskas (2010) ‘Measuring intrinsic quality of semantic search based on feature vectors’, takes component based analysis even further by conducting a sensitivity analysis on various components of their system. Their approach to semantic search is based on linguistic-semantic feature vectors generated from an ontology. They propose metrics to measure quality of feature vectors and run a set of experiments testing sensitivity of FVs with regards to tuning the system’s sub-components.

Accepted papers in this special issue cover the first two objectives of evaluation, without touching the third one. Typically, additional sophistication of the semantic systems adds complexity on user interaction to reach improved results. Consequently, standard IR metrics as recall and precision do not suffice alone to measure user satisfaction because of complexity and effort needed to use semantic search tools. But that is changing, to quote Lassila:

“After 10+ years of work into various aspects of the Semantic Web and its constituent technologies, I am now fully convinced (read: no longer in denial) that most of the remaining challenges to realise the Semantic Web vision have nothing to do with the underlying technologies involving data, ontologies, reasoning, etc. Instead, it all comes down to user interfaces and usability.” Lassila (2007)

We hope the readers will find the papers in this special issue interesting and useful in the application of their results. We also hope that these papers may inspire the readers’ own research on quality aspects of information retrieval systems.

References

- Blacoe, I.W., Tamma, V. and Wooldridge, M.J. (2010) ‘Evaluation of scalable multi-agent system architectures for searching the Semantic Web’, *Int. J. Metadata, Semantics and Ontologies*, Vol. 5, No. 2, pp.99–119.
- Lassila, O. (2007) Semantic Web Soul Searching. URL: http://www.lassila.org/blog/archive/2007/03/semantic_web_so_1.html [accessed: 2009/12/30].
- Tomassen, S.L. and Strasunskas, D. (2010) ‘Measuring intrinsic quality of semantic search based on Feature Vectors’, *Int. J. Metadata, Semantics and Ontologies*, Vol. 5, No. 2, pp.120–133.
- Uren, V., Sabou, M., Motta, E., Fernandez, M., Lopez, V. and Lei, Y. (2010) ‘Reflections on five years of evaluating semantic search systems’, *Int. J. Metadata, Semantics and Ontologies*, Vol. 5, No. 2, pp.87–98.