

Benchmarking Semantic Web Technology

Raúl García-Castro

Ontology Engineering Group,
Departamento de Inteligencia Artificial. Facultad de Informática,
Universidad Politécnica de Madrid, Spain
rgarcia@fi.upm.es

Abstract

This paper summarises the research problem that the PhD thesis I am currently writing addresses. It presents an overview of the thesis, its goals taking into account the deficiencies of the State of the Art, the approach followed, and the work performed for the thesis since its beginning in 2004.

1 Thesis overview

The Semantic Web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation [1]. The best way to introduce this new knowledge in the Web is through ontologies, and the Semantic Web needs technology both to create and to maintain these ontologies.

Ontology technology has improved considerably since its first tools were developed in the nineties. But to consolidate this technology both in the industrial and in the academic world, we need to evaluate it thoroughly to provide objective results. Up to now, the evaluation of this technology has seldom been carried out but, as its use spreads, numerous studies involving the evaluation of the Semantic Web technology have appeared, and several evaluation initiatives such as the *Evaluation of ontology-based tools* (EON) workshops or the *Ontology Alignment Evaluation Initiative* (OAEI)¹ have emerged in the last few years.

Benchmarking is a continuous process for improving products, services and processes by systematically evaluating and comparing them to those considered to be the best. This definition, adapted from the business management community [2], is used by some authors in the Software Engineering community [3] while others consider benchmarking as a software evaluation method [4].

The reason for benchmarking ontology tools instead of just evaluating them is to obtain several benefits from the former process that cannot be obtained from the latter. The evaluation of a tool shows us the weaknesses of that

¹<http://oaei.ontologymatching.org/>

tool or its compliance to quality requirements. If several tools are involved in the evaluation, we also obtain a comparative analysis of these tools and recommendations for users of these tools. When benchmarking several tools, besides all the benefits commented, we obtain a continuous improvement of the tools, recommendations for developers on the practices used when developing these tools and, from these practices, those that can be considered best practices.

On the other hand, to be able to facilitate and automate both the evaluation and the benchmarking tasks and to reach consensus in the field of ontology technology evaluation, it is necessary that both the research community and the industrial world have benchmark suites. Nowadays these benchmark suites are scarce and not general nor consensuated enough to be applied to a broad range of tools.

2 Thesis goals

The study of the State of the Art performed by the author [5] shows the following deficiencies when benchmarking the Semantic Web technology:

- The Semantic Web technology has not been evaluated either enough or objectively.
- No methodology exists for benchmarking the Semantic Web technology, though a methodology is necessary for extracting the best practices used when developing this technology, and for obtaining a continuous improvement.
- There are neither general sets of benchmark suites nor specific software for evaluating and benchmarking the Semantic Web technology.

The **methodological goals** of the doctoral thesis are the following:

- To develop a benchmarking methodology for the Semantic Web technology. This methodology will specify the tasks to be performed in the benchmarking process.
- To identify the techniques for performing each of the tasks of the proposed methodology.

The **technological goals** of the doctoral thesis are the following:

- To develop different benchmark suites for evaluating ontology development tools taking into account their performance, scalability, interoperability, etc.
- To develop tools for providing technological support to the different tasks of the benchmarking methodology.

3 Approach proposed

According to the thesis goals mentioned above, the tasks to perform in the thesis are the following:

1. **To investigate the State of the Art** in the subjects previously mentioned. This State of the Art includes an analysis of the existing definitions and methodologies for benchmarking, experimentation in Software Engineering, and Software Measurement. It also includes an analysis of the current Semantic Web technology evaluation initiatives.
2. **To develop a first version of a benchmarking methodology for the Semantic Web technology** from the analysis of the existing methodologies within the areas of benchmarking, of experimentation in Software Engineering, and of Software Measurement.
3. **To develop benchmark suites and tools for evaluating the performance and scalability of ontology development tools.** The benchmark suites and tools for evaluating the performance and scalability of ontology development tools will be defined in a generic way, so they can be used in several tools.
4. To validate the benchmarking methodology, the benchmark suites and the tools performing a **benchmarking of the performance and the scalability of ontology development tools.**
5. **To develop benchmark suites and tools for evaluating the interoperability of ontology development tools.** The benchmark suites and tools for evaluating the interoperability of ontology development tools will be defined in a generic way, so they can be used in several tools.
6. To validate the benchmarking methodology, the benchmark suites and the tools performing a **benchmarking of the interoperability of ontology development tools.**
7. **To improve the benchmark suites and the tools** developed in the previous tasks through the lessons learnt in the benchmarking activities.
8. **To propose a final version of the benchmarking methodology for the Semantic Web technology** through the lessons learnt in the benchmarking activities performed through the thesis.

4 Work performed

Most of the work performed in this thesis is being developed in the context of the *Knowledge Web* European Network of Excellence². In Knowledge Web,

²<http://knowledgeweb.semanticweb.org/>

benchmarking activities play a key role, being present both in the industrial and in the research fields.

The benchmarking of the performance and scalability of WebODE is being carried out in the CICYT project: *Infraestructura tecnológica de servicios semánticos para la web semántica*, with the goals of stating its precise performance and boosting the transference of the tool to industry.

From the tasks to perform in the thesis, which have been mentioned above, the following have already been performed or are in good progress:

- The State of the Art has been performed and is a chapter of the deliverable D2.1.1 of the Knowledge Web Network of Excellence [5].
- The benchmarking methodology has also been developed. It is described in the Knowledge Web deliverable D2.1.4 [6] and is being used in some benchmarking activities of Knowledge Web.
- In the context of the project *Infraestructura tecnológica de servicios semánticos para la web semántica*, a benchmark suite has been defined to evaluate the performance and scalability of WebODE. Also, different tools such as an ontology workload generator and an statistical data analyser for the evaluation results have been developed. With these benchmark suites and tools, WebODE has been evaluated and its actual performance and scalability have been determined. With the results of these tasks, several international papers have been published [7, 8].
- The benchmarking of the interoperability of ontology development tools using RDF(S) as interchange language is currently taking place³. Benchmark suites have been defined for evaluating the import and export from/to RDF(S) and the tools interoperability [9]. The author has organised the international benchmarking activity where different organisations participate with the most relevant ontology development tools. With the results of these tasks, several international papers have been published [10, 11].
- Also in Knowledge Web, benchmarking the interoperability of ontology development tools using OWL as interchange language has started⁴. The benchmark suites have been defined for evaluating the import and export from/to OWL and the interoperability of the tools.

5 Acknowledgments

This work has been supported by a FPI grant from the Spanish Ministry of Education (BES-2005-8024), by the IST project Knowledge Web (IST-2004-507482) and by the CICYT project *Infraestructura tecnológica de servicios semánticos para la web semántica* (TIN2004-02660).

³http://knowledgeweb.semanticweb.org/benchmarking_interoperability/

⁴http://knowledgeweb.semanticweb.org/benchmarking_interoperability/owl/

References

- [1] Tim Berners-Lee, James Hendler, and Ora Lassila. The Semantic Web: A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, May, 2001.
- [2] M.J. Spendolini. *The Benchmarking Book*. AMACOM, New York, NY, 1992.
- [3] C. Wohlin, A. Aurum, H. Petersson, F. Shull, and M. Ciolkowski. Software inspection benchmarking - a qualitative and quantitative comparative opportunity. In *Proceedings of 8th International Software Metrics Symposium*, pages 118–130, June 2002.
- [4] B. Kitchenham. DESMET: A method for evaluating software engineering methods and tools. Technical Report TR96-09, Department of Computer Science, University of Keele, Staffordshire, UK, 1996.
- [5] H. Wache, L. Serafini, and R. García-Castro. D2.1.1 survey of scalability techniques for reasoning with ontologies. Technical report, Knowledge Web, July 2004.
- [6] R. García-Castro, D. Maynard, H. Wache, D. Foxvog, and R. González-Cabero. D2.1.4 specification of a methodology, general criteria and benchmark suites for benchmarking ontology tools. Technical report, Knowledge Web, December 2004.
- [7] R. García-Castro and A. Gómez-Pérez. A benchmark suite for evaluating the performance of the WebODE ontology engineering platform. In *Proceedings of the 3rd International Workshop on Evaluation of Ontology-based Tools (EON2004)*, Hiroshima, Japan, November 8th 2004.
- [8] R. García-Castro and A. Gómez-Pérez. Guidelines for benchmarking the performance of ontology management APIs. In *Proceedings of the 4th International Semantic Web Conference (ISWC2005)*, number 3729 in LNCS, pages 277–292, Galway, Ireland, November 2005. Springer-Verlag.
- [9] R. García-Castro. D2.1.5 prototypes of tools and benchmark suites for benchmarking ontology building tools. Technical report, Knowledge Web, December 2005.
- [10] R. García-Castro and A. Gómez-Pérez. A method for performing an exhaustive evaluation of RDF(S) importers. In *Proceedings of the Workshop on Scalable Semantic Web Knowledge Based Systems (SSWS2005)*, number 3807 in LNCS, New York, USA, November 2005. Springer-Verlag.
- [11] R. García-Castro and A. Gómez-Pérez. Benchmark suites for improving the RDF(S) importers and exporters of ontology development tools. In *To appear in Proceedings of the 3rd European Semantic Web Conference (ESWC2006)*, Budva, Montenegro, June 2006.