

Renata Dividino, Massimo Romanelli and Daniel Sonntag

DFKI GmbH, Saarbrücken

renata.dividino@igd.fraunhofer.de, massimo.romanelli@dfki,de, daniel.sonntag@dfki.de

Introduction

Ontology evaluation is a basis for defining what a good ontology is.

By analyzing the heterogeneous nature of works with respect to the assessment of the ontology, we could note three groups of evaluation methods assessing:

S-OntoEval

Basically, the S-OntoEval's GUI is composed of three tab windows for the evaluation at three semiotic levels, respectively.

> Structural Functional Usability Structural Metrics - 📑 Topological .

The task-based evaluation approach is a gold standard-based method which involves the creation of a validated corpus of answers for a certain task. The gold standard is used as a reference for checking the performance of an ontology driven system.

At the usability level, S-OntoEval implements the annotation analysis metric.

- the graph structure and formal semantics of an ontology (Guarino and Welty, 2002; Yao et al., 2005; Huang and Diao, 2006);
- the ontology's intended use (Maedche and Staab, 2002; Porzel and Malaka, 2004; Daelemans and Reinberger, 2004; Lozano-Tello and Gomez-Perez, 2004); and
- the quality level of the ontology's annotations (Noy, 2004).

These three assessment approaches are directly analogous to a semiotic assessment.

Considering an ontology is a semiotic object (see figure below), the quality of an ontology may be assessed with respect to its graph (structure, syntax) and formal semantics), its **intended conceptualization** (cognitive semantics), and its **communication setting** (pragmatics).





Each tab presents specific UI components (e.g., trees) for the corresponding evaluation parameters.

Evaluation of SWIntO Use Case

We present the semiotic-based evaluation of the SWIntO ontology (Oberle et al., 2007) developed in the SmartWeb project (Wahlster, 2007).

SWIntO integrates a selection of specific domain and task specific ontologies (navigation, sport, discourse, etc.) with a core ontology.

We evaluated two metrics at the structural level:

- maximum depth: provides the number of nodes that lie on the longest path of the ontology's tree.
- consistency checking: checks logical adequacy of

Structural Annotations	Functional Annotations	🔄 Usability Annotations
Topological Depth Breath Fan-outness Size Logical Consistency Metalogical Meta-consistency	 Task Assessment Performance Pre/Pos Conditions I/O Pattern Topic Assessment Corpus Corpus Reusable Components 	Deployment Author Version Organization Content Annotation Content Annotation Price Disclaimer History Reviews
Total of annotated classes and properties: 1 Total of classes and properties: 21597 Class : http://smartweb.semanticweb.org/on Annotation: An Organism consisting of a core nucleic acid enclosed in a protective coat of p only inside a host living cell. A virus exhibits s usual characteristics of living things. Class : http://smartweb.semanticweb.org/on Annotation: The Class of Mammals with one c of incisors for gnawing. Includes rats, mice, p rabbits. Class : http://smartweb.semanticweb.org/on Annotation: Describes the physical format of Class : http://smartweb.semanticweb.org/on	tology/smartsumo#Virus of a single protein. A virus may replicate come but not all of the tology/smartsumo#Rodent ir two pairs guinea pigs, and tology/mpeg7#MediaInformation the multimedia data.	

The annotation analysis consists of quantifying the number of ontology elements linked to the tag rdf:comments.

Conclusion

We looked at existing ontology evaluation methods and implemented a semiotic based evaluation tool (S-OntoEval) from the perspective of their integration into a single framework. This is achieved by basically three steps:



Evaluation in Semiotic Levels

Our semiotic-based ontology evaluation tool (S-OntoEval) makes use of an unique evaluation framework (Dividino, 2007) which allows to assess the quality of ontologies by drawing upon semiotics.

It consists of three main modules (structural, functional and usability) each of them responsible for the assessment of the ontology in one semiotic dimension.



the ontology model, i.e., lack of contradiction of the ontology (none of the facts deducible from the model contradict one another).

The structural evaluation below follows the consistency checking approach.

				×
Structural Functional	Usability			
Structural Metrics	Metric	Туре	Evaluate	Weight
🕈 🗂 Topological	Depth	Topological		1
– 🗋 Depth	Breath	Topological		1
- 🗋 Breath	Fan-outness	Topological		1
- 🗍 Fan-outness	Size	Topological		1
	Consistency Moto, consistency	Logical		1
Size	Meta-consistency	Metalogical		1
P- ☐ Logical				Evaluate
P ☐ Metalogical	particular contraction of the second s			
Meta-consister	Ontology is Consistent			t-Box
Meta-consister	1			Thing
1000				- 🗋 Medialnformation
1000				- ContentAnswer
1000				=
1000				- D Medialdentification
100				— 🗋 ColorSpace
100				Ŷ
				👇 🚍 perdurant
200				👇 🚍 InternalChange
4144				- C Damaging
1000				
1000				
0000				
1000			•	🗕 🗋 Injuring
1000	General Structural	Functional		🕈 🗂 BiologicalProcess
1000	Structural			
	Choose one of the reaso	ners helour		🕈 🔚 AutonomicPr
1000	choose one of the reaso	ners below.		- 🗋 OrganOr1
				- Growth
				- 🔂 Breathing
1000				
	Reasoner: Racer		-	Digesting

The functional evaluation follows a **task-based** approach.

ructural Functional U	sability	
Functional Metrics Task Assessment Performance Pre/Pos Conditions I/O Pattern Topic Assessment Corpus Modularity Assessment Reusable Componem	welche Spiele fanden 1954 im Halbfinale statt wie spielte Deutschland gegen Brasilien wie hat Deutschland gegen Brasilien gespielt wer ist 2002 Weltmeister geworden wie ging das Finale 2002 aus wie ist das Finale 2002 ausgegangen wie spielte Brasilien 2002 wer war 1990 Weltmeister wer ist 1990 Weltmeister geworden wie snielte Deutschland hei der WM 1978	t-Box Thing Thing abstract ResultList MorphoSyntacticDecomposition Inflection Stemming

- firstly, the quality assessment of **ontology syntax** by making use of methods assessing the ontology's topological dimension and formal semantics (logical dimension);
- secondly, the **semantic dimension** by measuring the accuracy of the ontology with respect to its conceptualization (or intended use);
- and finally, its **usability dimension** by adopting approaches addressing the quality level of the set of annotations about the ontology and its elements.

In the evaluation example, we choose four metrics among others (depth, consistency checking, taskbased, and annotation analysis) which we believe to be representative in any ontology evaluation process.

The S-OntoEval tool allows to combine different evaluation scores which we plan to extend to more user-specific personalized combinations. After using S-OntoEval on the SWIntO ontology, an optimized version could be deployed for usage in the project THESEUS.

References

S-OntoEval makes use of a Semantic Web framework (Java/Jena), which provides a programming environment for parsing and interpreting RDF(S) and OWL documents.

Ontology Time Pe	is better perfomance than evaluated ontologer Prformance: 175	gy 🛔	t-Box
GS Ontology Time	e Performance: 97		Thing
emma#Emma Class http://smai al equivalent to h Hierarchy Overlap y/emma#Emma The classes http:	for class: http://smartweb.semanticweb.org rtweb.semanticweb.org/ontology/emma#Er ttp://smartweb.semanticweb.org/emma.ov os for class: http://smartweb.semanticweb. //smartweb.semanticweb.org/ontology/em web.semanticweb.org/emma.ovl#Emma.ov	nma is lexic vl#Emma org/ontolog ma#Emma	MediaInformation ContentAnswer MediaIdentification ColorSpace entity entity entity
% w.r.t their taxor		vertap in 50	P- 🗂 Damaging
			🕈 🗖 Destruction
	nomy	Open	← 🗖 Destruction
General Stru	nomy uctural Functional		 P-□ Destruction □ Killing □ Injuring P-□ BiologicalProcess P-□ PhysiologicProces
General Stru Gold Standard:	nomy uctural Functional /home/renata/root/GS/gs.owl	Open	P-□ Destruction ☐ Milling ☐ Injuring P-□ BiologicalProcess

The task-based approach measures the quality of the ontology with respect to its performance, given a particular task. It deals with measuring the quality of an ontology based on the adequacy of the ontological model of a specific task.

N. Guarino and C. Welty. 2002. Evaluating ontological decisions with Ontoclean. Commun. ACM, 45(2):61–65. Yao et al., 2005

N. Huang and S. Diao. 2006. Structure-based ontology evaluation. In ICEBE '06: Proceedings of the IEEE International Conference on e-Business Engineering, pages 132–137, Washington, DC, USA

A. Maedche and S. Staab, 2002. Measuring similarity between ontologies. Springer, Madrid, Spain

R. Porzel and R. Malaka. 2004. A task-based approach for ontology evaluation. In Proceedings of the ECAI 2004 Workshop on Ontology Learning and Population, Valencia, Spain, August

W. Daelemans and M. Reinberger. 2004. Shallow text understanding for ontology content evaluation. IEEE Intelligent Systems, 19(4):1541–1672

A. Lozano-Tello and A. Gomez-Perez. 2004. Ontometric: A method to choose the appropriate ontology. Journal of Database Management, 15:1–18

N. Noy. 2004. Evaluation by ontology consumers. IEEE Intelligent Systems, 19(4):74–81

W. Wahlster, 2007. 40 Jahre Informatikforschung in Deutschland, chapter SmartWeb: Ein multimodales Dialogsystem für das semantische Web. Springer, Heidelberg, Berlin

D. Oberle, et al. 2007. DOLCE ergo SUMO: On Foundational and Domain Models in the Smartweb Integrated Ontology (SWIntO). Web Semant., 5(3):156–174

R. Dividino. 2007. Semiotic-based ontology evaluation tool. Master's thesis, Universität des Saarlandes

Bundesministerium für Wirtschaft und Technologie

This research is supported by the German Federal Ministry of Economics and Technology under the grant number 01MQ07016 (THESEUS). The responsibility for this publication lies with the authors.