

Towards a Glossary of Activities in the Ontology Engineering Field

Mari Carmen Suárez-Figueroa, Asunción Gómez-Pérez

Ontology Engineering Group, Universidad Politécnica de Madrid

Campus de Montegancedo, Avda. Montepríncipe s/n, Boadilla del Monte, Madrid, Spain

E-mail: mcsuarez@fi.upm.es, asun@fi.upm.es

Abstract

The Semantic Web of the future will be characterized by using a very large number of ontologies embedded in ontology networks. It is important to provide strong methodological support for collaborative and context-sensitive development of networks of ontologies. This methodological support includes the identification and definition of which activities should be carried out when ontology networks are collaboratively built. In this paper we present the consensus reaching process followed within the NeOn consortium for the identification and definition of the activities involved in the ontology network development process. The consensus reaching process here presented produces as a result the NeOn Glossary of Activities. This work was conceived due to the lack of standardization in the Ontology Engineering terminology, which clearly contrasts with the Software Engineering field. Our future aim is to standardize the NeOn Glossary of Activities.

1. Introduction

The Semantic Web of the future will be characterized by using a very large number of ontologies embedded in ontology networks built by distributed teams (NeOn Consortium, 2006). So, future Semantic Web applications will be based on networks of contextualized ontologies, which will be in continuous evolution. With this new vision of ontologies and the Semantic Web, it is important to provide strong methodological support for collaborative and context-sensitive development of networks of ontologies in distributed environments. This methodological support, which is being created within the NeOn project¹, includes the identification and definition of activities that should be carried out when ontology networks are collaboratively built.

It was noticed that researchers, technology developers, and users used different terminology² to name the activities involved in the ontology development process. That is, no consensus had been reached yet on many of the definitions for ontology engineering activities. For instance, it was not clear enough the difference between ontology modification (Stojanovic, 2004) and ontology update (Stojanovic et al., 2002); and other activities had multiple definitions in natural language (e.g., ontology merging (Fernández-López et al., 1997; Kalfoglou & Schorlemmer, 2003; Kotis & Vouros, 2004)). Additionally, we can now observe that new activities related to the Semantic Web of the future are emerging without a concrete and precise definition (e.g. ontology modularization). This situation is the result of a lack of standardization in the Ontology Engineering terminology, which clearly contrasts with the Software Engineering field that boasts the *IEEE Standard Glossary of Software Engineering Terminology* (IEEE, 1990), a consensual glossary. We understand glossary as an alphabetical list of technical terms in some specialized

field of knowledge. In the case of *IEEE Standard Glossary of Software Engineering Terminology* (IEEE, 1990), this glossary identifies terms in use in the field of Software Engineering and establishes standard definitions for those terms.

Within the NeOn project, it has been decided to build the *NeOn Glossary of Activities* for unifying the terminology used by the NeOn partners. A unification of the set of terms, in our case activities, used within the NeOn project will enable a clearer communication between researchers, technology developers, and users participating in the project. The goal was to achieve consensus on the identification and unambiguous definition of the activities involved in developing ontology networks, which is the first attempt in the ontological engineering field. Out of the scope of this glossary are definitions of resources (data, metadata, etc.), which will be included in a Resource Glossary. The NeOn Glossary will include the NeOn Glossary of Activities, as an important and independent subset, and the Resource Glossary.

In order to reach a consensus in the Ontology Engineering terminology, focused on activities involved in the ontology development process, we decided to use the wiki technology (Leuf & Cunningham, 2001), which supports a higher level of consensus building by community members, because a user who disagrees with a statement can very easily modify it, delete it, comment it, etc. (Viégas et al., 2004). Wiki technology has already been used for obtaining consensus on ontology modelling during the ontology development (Hepp et al., 2007).

We created a non public space in the NeOn wiki for discussing the Ontology Engineering terminology, expressing and exchanging different opinions among different partners involved in NeOn, and reaching a final agreement. Meetings and mailing lists were also employed for agreeing on the activity definitions at final stages.

So, in this work we present the first step in an attempt to standardize the terminology used in the Ontology Engineering field. Such step lies in achieving consensus on the activities involved in the development process for

¹ <http://www.neon-project.org/>

² A system of words used to name things in a particular discipline.

ontology networks, within the NeOn consortium. From our understanding, any standardization agency such as ISO or W3C deals with the unification of Ontology Engineering terms. Only some ISO technical committees such as ISO/TC37/SC4 are working on the contribution of the ontologies for unifying linguistic resources.

The paper includes first the consensus reaching process for the NeOn Glossary of Activities and second the conclusion obtained.

2. Consensus Reaching Process for the NeOn Glossary of Activities

In this section we sketch the roles and the overall process gone through by the NeOn consortium to reach a consensus on the activities for developing ontology networks. During the process we have tried to achieve a consensus on the list of activities and on the activity definitions.

2.1. Roles in the Process

A varied number of skilled people, geographically dispersed (called ‘NeOn Glossary’ team), participated collaboratively in the consensus reaching process. The ‘NeOn Glossary’ team has a well-balanced and representative participation of people with different background: ontology engineers, ontology editors, and users within the NeOn project. The ‘NeOn Glossary’ team was composed of 25 people belonging to 9 institutions. The following concrete roles were distinguished:

- *The NeOn Glossary coordinator.* One of the 9 institutions was in charge of creating the NeOn wiki page dedicated to the consensus reaching process. This institution created and included in the wiki a template for gathering information about the activities identified in the literature for building ontologies. Additionally, the coordinator introduced for the debate the initial list of activities with initial definitions based on the study of the state of the art in Ontology Engineering.
- *The NeOn Glossary team.* The 9 institutions introduced definitions and commented them. They also participated in the different ad-hoc meetings carried out to reach consensus.

2.2. Process Stages

Before beginning with the consensus reaching process, the meanings of consensus and consensus reaching process were explained. The proposed process to achieve consensus on the activities was also explained and reviewed by the ‘NeOn Glossary’ team. After that, the team agreed on a targeted time period (one year) to reach consensus. And finally, the team followed the general process, shown in Figure 1, to achieve consensus on the activities involved in the development of ontology networks and their definitions. The consensus was achieved after the third round of the process.

For collecting the definitions in the NeOn Glossary, the ‘NeOn Glossary’ team based on three basic principles of

defining: “avoid circularity, define every term³ in a definition, and make sure that every term’s definition says what the term means”. In addition, the following principles were taken into account:

- Conciseness, i.e. every definition should say the most in the least number of words.
- Clarity in that it avoids ambiguity, i.e. words should be used unambiguously.
- Appropriateness, i.e. the definition should be appropriate to the target reader.
- Priority of essential traits, i.e. a definition should highlight the essential features of meaning.

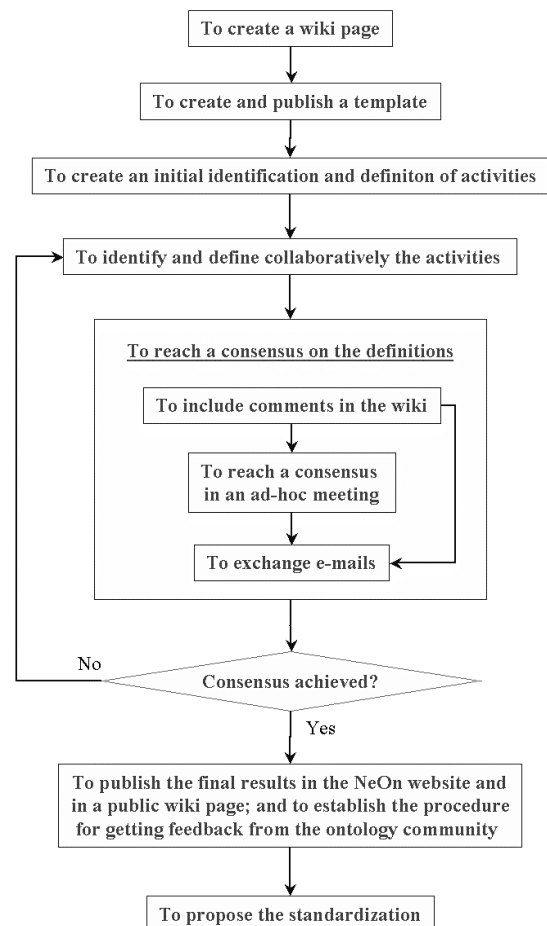


Figure 1: Consensus Reaching Process

The general process followed to achieve consensus on the activity terminology, shown in Figure 1, can be summarized as follows:

1. To create a NeOn wiki page dedicated to the consensus reaching process, within the NeOn consortium.
2. To create a template for gathering general information about activities (definition, classification, inputs, outputs, etc.) and to publish this template in the wiki. The template is shown in Table 1.

³ Terms may consist of simple words or complex phrases.

Template Slot	Description
<i>Activity Name</i>	Name of the activity
<i>Definition</i>	One or several natural language (NL) definitions (with the corresponding references)
<i>Type of Activity according to IEEE</i>	In the Software Engineering field, activities are grouped administratively into five main activity groups. Based on that, the following groups are proposed: (a) Ontology Management Activity (b) Ontology Pre-Development Activity (c) Ontology Development Activity (d) Ontology Post-Development Activity (e) Ontology Support Activity The type for a concrete activity should be unique
<i>Input</i>	A list of the required information to be input of the activity
<i>Output</i>	A list of the information that is required to be output of the activity
<i>References</i>	References for the NL definitions
<i>Comments</i>	Other comments about the activity

Table 1: Template for the Activities in the NeOn Glossary

3. To create an initial NeOn Glossary of Activities, including initial identification and definition of the main activities to be included in the ontology network development process.

For creating this initial glossary, the NeOn Glossary coordinator manually extracted terms (in our case, activities for building ontologies) and their definitions from books, papers and documents well-known in the Ontology Engineering field.

This initial glossary was made available on the wiki following the template presented in Table 1.

4. To identify and define activities collaboratively in the wiki according to the initial glossary. People in the ‘NeOn Glossary’ team were totally free to incorporate more activities and/or definitions in the initial glossary; and to include more general information about the activity (such as, input and output, classification of the activity following the groups based on IEEE, etc.).

5. To reach consensus on the activity definitions. In this stage the ‘NeOn Glossary’ team used the wiki, ad-hoc meetings and e-mails, and the ‘NeOn Glossary’ team adopted the following process:

- a. All institutions participating in the ‘NeOn Glossary’ team commented the majority of the activities in the wiki.
- b. Two ad-hoc meetings were needed for reviewing and agreeing on definitions of activities. The rules for accepting, or not, a concrete activity definition were:
 - If the team’s comments were generally positive and no major objections were raised, then the definition was considered as final.

- If general comments were positive, but someone had a major objection to the definition, the definition was modified until no major objection was encountered.
- If the team’s comments were generally negative, the definition was ruled out.
- If the team’s comments were mixed, there were three possibilities:
 - discussions continued until positive or negative results were achieved;
 - discussions were postponed until the next meeting; and
 - the issue was postponed until more information was available in the wiki.
- If discussions seemed to be going on forever without the possibility of reaching an agreement, the team could:
 - decide to drop the definition, or the activity; or
 - move onto approval by voting the definition. The selected voting procedure was based on absolute majority.

c. To exchange e-mails with the rest of the NeOn consortium members not directly involved in the glossary creation with the goal of getting a final agreement at the consortium level.

6. To publish the final results on the NeOn website⁴ and on a public wiki page and to establish the procedure for getting feedback from the Ontology Engineering community, using the argumentation tool ‘Cicero’⁵ (which is already a wiki).

The current version of the NeOn Glossary of Activities includes definitions of 53 activities, which are shown in Table 2. The current activity definitions can be downloaded from a dedicated web page in the NeOn website⁶.

Entries in the NeOn Glossary are arranged alphabetically. Additionally, notes have been added to clarify similar activities, and information about synonymous activities has been also included.

The vocabulary included in the glossary is monolingual (English).

7. To propose the standardization of the NeOn Glossary of Activities.

The idea is to propose to standardization committees, such as the technical committee ISO/TC37, the standardization of the NeOn Glossary.

Terminology standards help to avoid confusion by harmonizing terms, in our case activities involved in the development of ontology networks.

The future standard NeOn Glossary of Activities is intended to serve as useful reference for those in the Ontology Engineering field and for those who come into contact with ontologies.

⁴ <http://www.neon-project.org/>

⁵ <http://cicero.uni-koblenz.de>

⁶ <http://www.neon-project.org/web-content/images/Publications/neonglossaryofactivities.pdf>

<ul style="list-style-type: none"> • O. Aligning • O. Annotation • O. Assessment • O. Comparison • O. Conceptualization • O. Configuration Management • Control • O. Customization • O. Diagnosis • O. Documentation, • O. Elicitation • O. Enrichment • O. Environment Study • O. Evaluation • O. Evolution • O. Extension • O. Feasibility Study • O. Formalization • O. Forward Engineering • O. Implementation • O. Integration • Knowledge Acquisition for Ontologies • O. Learning • O. Localization • O. Mapping • O. Matching • O. Merging 	<ul style="list-style-type: none"> • O. Modification • O. Modularization • O. Module Extraction • O. Partitioning • O. Population • O. Pruning • O. Quality Assurance • Non Ontological Resource Reengineering • O. Reengineering • O. Restructuring • O. Repair • Non Ontological Resource Reuse • O. Reuse • O. Reverse Engineering • Scheduling • O. Search • O. Selection • O. Specialization • O. Specification • O. Summarization • O. Translation • O. Update • O. Upgrade • O. Validation • O. Verification • O. Versioning
---	---

Table 2: List of the Activities in the NeOn Glossary

3. Conclusion

The consensus reaching process for building the NeOn Glossary of Activities has been presented in this paper. The NeOn Glossary of Activities, which identifies and defines the activities potentially involved in the ontology network construction, is a first step for solving the lack of a standard glossary in the Ontology Engineering field in contrast with the *IEEE Standard Glossary of Software Engineering Terminology* (IEEE, 1990) in the Software Engineering field.

We already published the NeOn Glossary of Activities in the NeOn website⁷ and we still delve into the idea of obtaining feedback from the Ontology Engineering community (outside NeOn). In fact, we are creating a public wiki page with all the activities in the glossary to obtain comments from other people. We are using the argumentation tool “Cicero”⁸, which is already a wiki, so that the Ontology Engineering community will have the opportunity to comment the activity definitions for about one year. The long term goal is to have a more complete and consensual glossary, which could become the terminological reference in the Ontology Engineering field.

Finally, if the Ontology Engineering community supports the activity of providing feedback on the NeOn Glossary

of Activities, then we could think of approaching standardization agencies such as IEEE, ISO or W3C for the standardization of the NeOn Glossary of Activities.

4. Acknowledgements

This work has been supported by the NeOn project (IST-2005-027595). We are very grateful to our NeOn partners for their revisions and comments. We are also very grateful to Rosario Plaza, Guadalupe Aguado de Cea, and Elena Montiel-Ponsoda for their revisions and comments.

5. References

- IEEE Std 610.12-1990. *IEEE Standard Glossary of Software Engineering Terminology*.
- NeOn Consortium. NeOn: Lifecycle Support for Networked Ontologies. *NeOn Technical Annex*. 2006.
- Fernández-López, M., Gómez-Pérez, A., Juristo, N. (1997). *METHONTOLOGY: From Ontological Art Towards Ontological Engineering*. Spring Symposium on Ontological Engineering of AAAI. Stanford University, California, pp 33–40.
- Hepp, M., Siorpaes, K., Bachlechner, D. (2007). *Harvesting Wiki Consensus: Using Wikipedia Entries as Vocabulary for Knowledge Management*. IEEE Internet Computing 11(5): 54-65.
- Kalfoglou, Y., Schorlemmer, M. (2003). *Ontology mapping: the state of the art*. The Knowledge Engineering Review. Volume 18, Issue 1. PP: 1-31. ISSN:0269-8889.
- Kotis, K., Vouros, G. (2004). *The HCONE Approach to Ontology Merging*. In: The Semantic Web: Research and Applications. First European Semantic Web Symposium (ESWS). PP: 137-151.
- Leuf, B., Cunningham, W. (2001). *The Wiki Way*. Addison-Wesley.
- Stojanovic, L. (2004). *Methods and Tools for Ontology Evolution*. Dissertation. Referee: Rudi Studer, Christof Weinhardt, Asunción Gómez-Pérez.
- Stojanovic, L., Stojanovic, N., Handschuh, S. (2002). *Evolution of the Metadata in the Ontology-based Knowledge Management Systems*. 1st German Workshop on Experience Management: Sharing Experiences about the Sharing of Experience. Berlin.
- Viégas, F. B., Wattenberg, M., Dave, K. (2004). *Studying cooperation and conflict between authors with history flow visualizations*. Conference on Human Factors in Computing Systems. Vienna, Austria. Pages. 575-582. ISBN:1-58113-702-8.

⁷ <http://www.neon-project.org/>

⁸ http://cicero.uni-koblenz.de/wiki/index.php/Main_Page