



AsIsKnown

**A semantic-based knowledge flow system for
the European home textiles industry**

Work package 3: Common sense ontology engineering

**Deliverable D11 “Reports on the tests and evaluation of common
sense ontology”**

Lead participant: IPP-BAS
Nature: Report
Dissemination level: PU
Delivery date: 11 PM



This document has been produced in the context of the AsIsKnown Project. The AsIsKnown project is part of the European Community's Sixth Framework Program for research and development and is as such funded by the European Commission. All information in this document is provided "as is" and no guarantee or warranty is given that the information is fit for any particular purpose. The user thereof uses the information at its sole risk and liability. For the avoidance of all doubts, the European Commission has no liability in respect of this document, which is merely representing the authors view.



Context

WP 3	Common sense ontology engineering
Task 3.7	Testing and evaluation
Dependencies	This deliverable requires user requirement input from 3.2 and 3.4

Contributors:	Reviewers:
Kiril Simov (IPP-BAS) Petya Osenova (IPP-BAS)	Kiril Simov (IPP-BAS)

Approved by: Kiril Simov, Bulgaria as WP3 Head



Executive Summary

This report focuses on Tests and Evaluation of Common Sense Ontology. First, the general approaches to the problem are considered. Then, these approaches are parameterized with respect to the Home Textile area within AsIsKnown project.

Although being an active research topic recently, the testing and evaluation of ontologies is still far from standardization. Hence, there is no 'one agreed upon' methodology. Instead, various criteria from different points of view have been proposed within the relevant literature. To sum up, one stream of works on evaluation reflects the ontology performance as a whole, or as divided in its components. Another stream of researchers views the ontology evaluation as consisting of three dimensions: structural, functional and usability-oriented. The last, but not least, stream connects the evolution of the ontology with its evaluation.

As a main support component for knowledge transfer and translation among agents, the AsIsKnown ontology will be responsible for the domain data management and exploitation. Also, it has to be helpful for both target groups - domain experts and end users. For that reason its validation has become a very important step.

The ontology creation within the project is data-driven. It relies on standards and terminological lexicons. We envisage two types of validity checks. The first will come through the connection of the domain ontology to the upper ontology (DOLCE), and the second – through validation by domain experts. This validation will be performed by the means of specially designed questionnaires. They will reflect the current organization of the ontology knowledge. Hence, we view this kind of validation as a dynamic one. The pursued tasks with respect to the ontology are as follows: (1) extending the list of terms and relations, (2) correcting the inconsistencies, and (3) smoothing the gap between the domain knowledge as represented in the official document repositories (standards and lexicons), and as considered by the domain experts in their actual work.



Table of Contents

EXECUTIVE SUMMARY	4
TABLE OF CONTENTS	5
TABLE OF FIGURESFEHLER! TEXTMARKE NICHT DEFINIERT.	
LIST OF ABBREVIATIONS	6
1 INTRODUCTION AND PROBLEM STATEMENT	7
2 METHODOLOGIES FOR ONTOLOGY EVALUATION	8
3 ONTOLOGY EVALUATION IN ASISKOWN	11
4 CONCLUSIONS AND OUTLOOK	14
REFERENCES	15



List of Abbreviations

DOLCE	Descriptive Ontology for Linguistic and Cognitive Engineering
OntoWordNet	WordNet 1.6 aligned to DOLCE
OWL	Web Ontology Language
OWL-DL	Web Ontology Language – Description Logic sublanguage



1 Introduction and Problem Statement

In this deliverable we discuss the test and evaluation of the common sense domain ontology in the area of Home textile. In our work we accept the ideas of [3] that evaluation of an ontology is part of the life cycle of its development, and therefore it takes place in parallel to this development. Consequently, evaluation becomes a necessary procedure after each major step in the ontology creation and management.

The question of ontology testing and evaluation has been explored very much by researchers recently. However, there is no one unified approach to follow. The evaluation strategies and appropriateness became dependant on the specific tasks and aims. For that reason, in the report we first discuss various methodologies to ontology validation, and then, we present our approaches tuned to the requirements of the AsIsKnown project.

We support the opinion in [5] according to which 'Good ontologies are the ones that serve their purpose'. In order to do so, ontologies have to convey the appropriate piece of knowledge in the appropriate way and at the appropriate time. It would be difficult to evaluate the quality of such a knowledge transition due to the fact that knowledge has always a subjective interpretation. For that reason we prefer to adopt the incremental evaluation approach instead of having a separate evaluation level. Thus we could determine the problematic points regularly and adjust the structure of the ontology conceptualization accordingly.

The main milestones with respect to the incremental ontology evaluation seem to be the following: (1) ontology creation and (2) ontology parameters. The first milestone concerns the reliability of the knowledge sources, from which the ontology is derived. In our case these are international standards in the area as well as terminological lexicons. The second milestone has to do with ontology function, ontology formal representation, ontology coverage, and degree of communication with the other components within the general infrastructure. Ontology function in AsIsKnown architecture is manifold. The ontology has to serve as a 'knowledge translator' between humans and humans; humans and computer components; computer components and computer components. Concerning the second parameter, we rely on OWL-DL as language, which is flexible enough to cover the above mentioned requirements. Ontology coverage depends very much on the underlying sources. The communication with other components is not a trivial parameter, because we have to take into account the type of the input knowledge, and to communicate comprehensively the ontology knowledge further.

At this stage of our work we concentrate on two main tests on ontology quality: (1) checks over trend texts and (2) consultations with domain experts. The first check will test the coverage of the ontology. The second one is oriented towards the precision of the relations within and among concepts. Both tests will be validated by the connection to the upper ontology.

2 Methodologies for Ontology Evaluation

Evaluation of ontologies is an active research area. Although there already exist many proposals for evaluation of different aspects of the ontology development, usage, evolving, there is no a widely accepted evaluation methodology. In this section we present a short overview of the existing methodologies. In the next section we present our plan for evaluation of the common sense ontology developed within AsIsKnown project.

The overview presented here is based on [1], [2], [3] and [4]. The first classification (see [1]) of the methodologies for evaluation of ontologies is “*black box*” *evaluation methodologies* which are oriented towards testing the performance of tasks of the whole system, and “*glass box*” *evaluation methodologies* which are oriented towards testing different components of a system. The *black box evaluation* depends on the participation of human user who performs the corresponding task as an end user of the system. Thus the evaluation becomes very much application dependant and hence needs specific evaluation criteria. The *glass box evaluation* is additionally divided into three stages: (i) evaluation of an ontology in its pre-modelling stage, (ii) evaluation of an ontology in its modelling stage, and (iii) evaluation of an ontology after its release. The evaluation at stage 1 is in fact an evaluation of materials available to the ontology engineer when he or she starts the creation of the ontology. Such materials include other ontologies to be re-used in the creation of the new one, database schemata, domain oriented text corpora, etc. Before the actual usage of these sources the ontology engineer needs to evaluate their quality and appropriateness with respect to the creation of the new ontology. The evaluation at stage 2 comprises regular checks on the consistency of the ontology, logical errors or methodological errors with respect to the ontology construction methodology. At this stage also the ontology content is evaluated – whether the domain model represents the necessary concepts. The evaluation at stage 3 can involve comparison of the ontology to other knowledge sources (other ontologies or other conceptualizations of the same domain). Usually this evaluation is performed by people other than the ontology developers.

According to [4] there are three dimensions to which ontology can be evaluated: the *structural* dimension which concentrates on syntax and the formal semantics of the ontology (representation as a graph). Here the topological, logical and meta-logical properties of an ontology can be measured; the *functional* dimension reflects the intended usage of the ontology and its components; the *usability-profiling* dimension focuses on the communication context of an ontology. The structural measurements include the following among others: depth of the graph, breadth of the graph, tangledness, consistency, etc. The functional measurements refer to the correspondence between the ontology as a model, and the some conceptualization of the domain. The metrics, which are usually used in the evaluation of systems in Information Retrieval or NLP, such as precision, recall and accuracy are not straightforward applicable to ontology evaluation (see [5]), because it is not clear what exactly the ontology is compared to. However, [4] defines the following sources of information that can be used: (1) experts’ judgment and (2) a data set assumed as a qualified expression of experts’ judgment:

- *Agreement*: it is measured through the proportion of agreement that experts have with respect to ontology elements; when a group of experts is considered, we may also want to measure the consensus reached by the group’s members.
- *User-satisfaction*: it can be measured by means of dedicated polls, or by means of provenance, popularity, and trust assessment.

- *Task*: what has to be supported by an ontology? It deals with measuring of the ontology according to its conformity to some goals, preconditions, postconditions, constraints, options, etc. This makes the measurement very reliable at design-time, while it needs a reassessment at reuse-time.
- *Topic*: what are the boundaries of the knowledge domain addressed by an ontology? It deals with measuring the ontology according to its conformity to an existing knowledge repository. This makes the measurement reliable both at design-time, and at reuse-time, but is based on the availability of data that can be safely assumed as related to the (supposed) topic, covered by this ontology. Natural Language Processing (NLP)-based methods fit into this category, and are currently the most reliable methods for ontology evaluation, at least for lightweight ontologies.
- *Modularity*: what are the building blocks for the design of an ontology? It is based on the availability of data about the design of the ontology. Therefore, it deals with measuring of the ontology according to its conformity to an existing repository of reusable components. This makes the measurement very reliable both at design-time, and at reuse-time. On the other hand, modularity assessment is only practicable on ontologies that have been designed with the appropriate methodology.

The usability-profiling comprises three basic levels of characterisation: the *recognition* level refers to the instructions for using of the ontology (how easy it is to understand the ontology); the *efficiency* level is related to organisational, commercial and development characteristics of the ontology; the *interfacing* level reflect the process of matching the ontology to a user interface. In order to formalise these dimensions of ontology evaluation [4] defines a meta ontology O^2 which characterises ontologies as semiotic objects. We will not outline the ontology here, but we will repeat the list of quality-oriented ontology descriptions (principles) given by [4] as emerged in the practice of ontology engineering:

- *Cognitive ergonomics*. This principle prospects an ontology that can be easily understood, manipulated, and exploited by final users.
- *Transparency* (explicitness of organizing principles). This principle prospects an ontology that can be analyzed in detail, with a rich formalization of conceptual choices and motivations.
- *Computational integrity and efficiency*. This principle prospects an ontology that can be successfully/easily processed by a reasoner (inference engine, classifier, etc.).
- *Meta-level integrity*. This principle prospects an ontology, which respects certain ordering criteria, assumed as quality indicators.
- *Flexibility* (context-boundedness). This principle prospects an ontology that can be easily adapted to multiple views.
- *Compliance to expertise*. This principle prospects an ontology that is compliant to one or more users' knowledge.
- *Compliance to procedures for extension, integration, adaptation*, etc. This principle prospects an ontology that can be easily understood and manipulated for reuse and adaptation.
- *Generic accessibility* (computational as well as commercial). This principle prospects an ontology that can be easily accessed for effective application.
- *Organizational fitness*. This principle prospects an ontology that can be easily deployed within an organization, and that has a good coverage for that context.

The last point we would like to mention here is the approach taken by [3] which interconnects the evolution of an ontology and its evaluation. The authors define the quality of an ontology as a function of the context of the ontology and then the evolution is considered as an optimization procedure that maximizes the function. Although in our work we will not define such a function we will use their idea of interconnecting the development of the ontology and the regular evaluation of the intermediate states of the ontology. With respect to above criteria for the quality of an ontology we will try to reflect all of them in the process of development of our ontology.

3 Ontology Evaluation in AsIsKnown

As it was described in Deliverable D8, the AsIsKnown Home Textile Ontology defines the common language for representation of product data, ordering data, trend descriptions, and interior designs. In the first place the ontology has to provide a description of the features of the products from the user point of view. This will facilitate the arrangement of the products in product sets and interior designs following the casual user who is not an expert in the textile. On the other hand, the ontology needs to reflect the different physical and chemical properties of the materials used in producing the textile in order to satisfy the safety requirements imposed by the standards in a given country. Thus, our goal is to construct an ontology with a *wide coverage* of the domain of home textile, which is also *lexicalized* in several languages (initially we envisage lexicons in English, German, French and Bulgarian to be aligned to the ontology, but later the consortium decided we to cover all project language including Hungarian and Italian). The usage of the ontology within the AsIsKnown system will be mediated by the Smart Profiler which will play the role of filter between the user and the ontology. Thus the user will see as much from the ontology as it is necessary for the completion of a particular task.

We agree with [3] that evaluation is part of life cycle of the development of the ontology and we will evaluate the common sense ontology in home textile industry at different moments of its development and usage. Thus, in order to guarantee the quality of the ontology we accepted an ontology development methodology which incorporates steps that satisfy the principles given above. Here we present again the steps of our methodology and explain how each step contributes to the satisfaction of the quality descriptions.

As it is described in Deliverable 8, AsIsKnown ontology development methodology comprises the following steps:

- **Processing of the standards and vocabularies in the domain**

The development of the AsIsKnown Home Textile Ontology starts with processing of standards and terminological lexicons in the domain of textile and home textile (a list of the available standards is given in Deliverable D8). Being created by leading experts in the domain with the goal to facilitate the whole process of production and usage of the home textile, the standards can be viewed as “expert questionnaires” usually used in the process of knowledge acquisition. Thus, we expect to find definitions of the most important concepts and relations in the domain of textile. The definitions also will help us to establish the main relationships between the extracted concepts. As a means for the extraction of the concepts and the relations we are use partial analysis of the definitions. Then we inspect manually the analysis in order to identify the relevant knowledge. The result of this step will be a list of (concept) terms (in English), a list of relations (relational terms), a list of triples – (term1 relation term2), and additional information, such as the number of some objects that are related to an instance of a given concept. These lists will be the backbone of the ontology.

The selection of standards and terminological lexicons as a starting point for elicitation of the domain knowledge plays manifold role with respect to the evaluation of the ontology. First, it ensures a wide coverage of the ontology with respect to the domain. We expect most of the important concepts and relations in the domain to be defined and used in the standards and the terminological lexicons. Second, each concept (relation) defined in the ontology (or almost each of them) will be connected to particular definitions or explanations in the standards, thus, it will be easy to explain to the potential users the meaning of the concepts (relations) in the ontology. Third, the exploitation of such sources will minimize the cost of the development of the ontology, because the expensive participation of domain experts will be postponed to a later stage of the development of the ontology.

- **Formalisation of the terms**

The next step will be to define formal definitions of the extracted concepts and relations in OWL-DL – see [6]. We have selected OWL-DL, because there exist implemented reasoners for it. For each term in the term list we construct a class definition in OWL-DL. We do the same for each relational term. We also encode the additional information in the definitions of the terms and the relations. The result of this step is an initial formal version of the ontology. It is possible that the level of errors in the concepts and the definitions will be high. The possible errors are of two kinds – *ontological* and *domain-related*. The first ones are connected with the evaluation of the defined concepts and the relations between them with respect to meta-ontological properties. For the purpose to clean these errors we are going to use the OntoClean approach – see [7], [8] and [9]. The second kind of errors will be repaired with the help of the partners who are experts in the area of textile – see also below.

Encoding of the ontology at this stage of its development in a formal language which allows automatic inference is a prerequisite for regular consistency checks which would facilitate an early discovery of logical inconsistency in the ontology definitions.

- **Link to an upper ontology**

The establishing of the connection between the upper and the domain ontology will help us to check the consistency of the domain ontology with respect to the ontology construction methodology behind the upper ontology and to inherit the knowledge encoded in the upper ontology to the domain ontology. We have selected DOLCE Ontology as our upper ontology.

DOLCE ontology is descriptive, multiplicative ontology which adopts possibilism as an approach to existence and perdurantism as an approach to change. Selecting DOLCE as an upper ontology for our domain ontology we also adopt the formal background on which this ontology is constructed – OntoClean approach. OntoClean approach assigns meta-properties to the concepts in the ontology and facilitates the correctness of the taxonomic relations. All the concepts defined in DOLCE are annotated with the meta-properties of OntoClean. Via the mapping of the domain ontology to DOLCE we ensure inheritance of the meta-properties from DOLCE to the domain ontology and in this way it will facilitate the annotation of the domain concepts with the OntoClean meta-properties. The mapping to DOLCE is done via OntoWordNet [10].

- **Evaluation by domain experts**

The evaluation of the first version of the ontology will be done in two ways:

Practical evaluation

The ontology will be incorporated within AsIsKnown system. This task will be performed in the second year of the project.

Expert evaluation

Then the ontology will be reviewed by the partners in the project. The review will be done on the basis of manual examination of portion of the ontology and by filling of specially designed questionnaires. The questionnaires will be orientated to examine problematic places in the ontology. Such places are due to the fact that there was no enough information about a term in the textual materials on which the ontology is based, or the information is somewhat contradictory. The result of the evaluation will be in the form of requirements for introducing of new concepts, new relations or restructuring of the ontology. In the second turn of the development of the ontology we will incorporate these requirements. The resulting ontology will be the second version of the ontology.

There are several types of questions:

1. From a structural point of view: yes/no questions; open questions, which require a free answer; multiple choice questions.
2. From a content point of view: checking the validity of a specific relation between concepts; and providing a definition, when it is missing or is incomplete.

Here we give some sample questions from such a questionnaire. The questions are constructed on the basis of the concept definitions in the ontology. The answers from the domain experts will help us to figure out some tacit information within the standards and also to improve the structural properties of the ontology. Examples:

Are the terms **Flocked Pile Textile Floor Covering** and **Flocked carpet** the same thing?

The goal of such a question is to determine equivalence of two terms defined in the standards or terminological lexicons for which it is not clear whether they denote the same concept or there are different concepts for each term.

List the properties which differentiate between **Carpet** and **Mat**.

Two terms that are accepted to denote different concepts do not have defined (in the ontology) properties that discriminate between them.

Does **Batt** consist of **Fibres**?

This is a question about relations between two concepts.

After the generation of the questionnaire it is answered by several domain experts and the new information is incorporated in the ontology. We envisage this process to become (semi)-automatic.

- **Documentation**

In the process of construction of the ontology we keep track on the sources of each concept, relation or axiom. The track records point to the context from where the concept originated. Also, we keep information about all changes in the definition of the concept or the relation. Important part of the documentation will be the definitions in natural language created during the initial stage of concept definitions. This point is in direct connection to usability of the ontology.

The current state of the development of the AsIsKnown Home Textile Ontology allows us to do evaluation on the coverage and consistency of the ontology. The actual usability evaluation will be done during the second and the third year when the ontology is incorporated within AsIsKnown system.

Our impression is that the knowledge presented in the lexicons and the standards is not systematically organized. Thus, it requires reconstruction with respect to various relations (is-a, made-of, used-for etc.). At the moment only one or few of them are given priority while the others are suppressed.

4 Conclusions and Outlook

In this report we present a short overview of evaluation methodologies described in literature. On the basis of this we selected several principles of ontology qualities that have to be met during the development of AsIsKnown Home Textile Ontology.

Then we have presented how these principles are incorporated within the ontology development methodology adopted for AsIsKnown project. We have shown that most of the tests are incorporated in the methodology and thus they are satisfied during the creation of the ontology.

Some of the required checks are related to the further usage of the ontology within the project and they will be performed in the next project phases.

References

- [1] Jens Hartmann, Peter Spyns, Alain Giboin, Diana Maynard, Roberta Cuel, Mari Carmen Suárez-Figueroa and York Sure. 2005. *D1.2.3 Methods for ontology evaluation*. Deliverable D1.2.3 EU-IST Network of Excellence IST-2004-507482 KEWB.
- [2] Janez Brank, Marko Grobelnik, Dunja Mladenić. 2005. *D1.6.1 Ontology Evaluation*. Deliverable D1.6.1 EU-IST Project IST-2003-506826 SEKT.
- [3] Peter Haase, York Sure. 2005. *D3.1.2 Incremental Ontology Evolution-Evaluation*. Deliverable D3.1.2 EU-IST Project IST-2003-506826 SEKT.
- [4] Aldo Gangemi, Carola Catenacci, Massimiliano Ciaramita, and Jos Lehmann. 2006. *Modelling Ontology Evaluation and Validation*. To appear in Proceedings of ESWC2006, Springer.
- [5] Christopher Brewster, Harith Alani, Srinandan Dasmahapatra, and Yorick Wilks. 2004. *Data Driven Ontology Evaluation*. International Conference on Language Resources and Evaluation, Lisbon, Portugal.
- [6] OWL. *Web Ontology Language (Overview)*. <http://www.w3.org/TR/owl-features/>
- [7] Aldo Gangemi, Nicola Guarino, Claudio Masolo and Alessandro Oltramari. 2001. *Understanding top-level ontological distinctions*. Proc. of IJCAI 2001 workshop on Ontologies and Information Sharing.
- [8] Christopher Welty and Nicola Guarino. 2001. *Supporting Ontological Analysis of Taxonomic Relationships*. Data and Knowledge Engineering.
- [9] Nicola Guarino and Christopher Welty. 2002. "Evaluating Ontological Decisions with *OntoClean*." Communications of the ACM, 45(2): 61-65.
- [10] Aldo Gangemi, Roberto Navigli, and Paola Velardi. 2003. *The OntoWordNet Project: extension and axiomatisation of conceptual relations in WordNet*. International Conference on Ontologies, Databases and Applications of SEMantics (ODBASE 2003), Catania, Italy.