# A Semantic-based Approach for Ontology Module Extraction

Amir Souissi[1], Walid Chainbi[2] and Khaled Ghedira[3]
[1]*National School of Computer Science, SOIE, Manouba University, Manouba, Tunisia*
[2]*Sousse National School of Engineers, SOIE, Sousse University, Sousse, Tunisia*
[3]*Higher Institute of Management, SOIE, Tunis University, Tunis, Tunisia*

Keywords: Ontologies, Modularization, Extraction, Semantics.

Abstract: Ontology modularization is crucial to support knowledge reuse on the ever increasing Semantic Web. However, modularization methods that serve the reuse goal are often intended for humans to assist them in building new ontologies, rather than for applications that need only a relevant part of an existing ontology. Moreover, modules obtained are always subject to verification and maintenance by humans to validate the semantic consistency of their contents. In this paper, we investigate how semantic comparisons may provide a module relevant to a set of terms which are not part of the ontology. Our objective is to extract a module which may be usable as a separate ontology. The user does not need to be familiar with the exact terms used inside the ontology beforehand to extract from it a module for a specific application/knowledge sub domain.

## 1 INTRODUCTION

Ontologies have established themselves as a powerful tool to enable knowledge sharing, and a growing number of applications have benefited from the use of ontologies as a means to achieve semantic interoperability among heterogeneous, distributed systems. Ontologies play a key role in one of the newest areas of interest, the Semantic Web, as confirmed by efforts such as OntoWeb, and OWL. A widely quoted definition of an ontology was proposed by Gruber who defines it as an explicit specification of a conceptualization (Gruber, 1993). An ontology specifies a vocabulary including the key terms, their semantic interconnections, and some rules of inference.

With the evolution of cooperative and distributed systems, and the emergence of the Semantic Web, ontologies have become an indispensable resource. The number of ontologies available on the Web has also increased due to the appearance of several tools that assist users in creating their ontologies. This has posed problems of understanding and reuse of those resources already difficult to design. A solution was then proposed by knowledge engineers namely *modularization*. Spaccapietra indicates in (Spaccapietra, 2005) that a module is a subset of a whole that makes sense and can somehow exist separated from the original ontology and not necessarily supporting the same functionality as it. He highlights five goals to modularization which are scalability, complexity management, understandability, personalization and reuse. He considers that the understanding of what modularization exactly means and what are the advantages and the disadvantages which are expected from modularization depend on these goals assigned to modularization. Since ontology construction is a labor intensive task and it is time consuming, the modularization methods which serve the purpose of reuse often focus at reusing ontological modules for building new ontologies (Cuenca Grau et al., 2007b; Doran et al., 2007). The focus of this paper is on ontology modularization for reuse. We aim to extract a part from an ontology in a way such that it can be reused as an ontology instead of the original one. Our objective is to allow obtaining a module which covers a specific topic from the ontology and to consider this module as a new ontology modeling this topic. Since some current ontologies are evolving to more expressivity and complexity, we propose an approach which targets ontologies without a clear internal structure (more semantic relations and hierarchical staple relations). Our approach intends to extract a module relevant to a set of terms which may be different of these employed inside the ontology. The idea is to extract a module without being necessarily familiar


Souissi A., Chainbi W. and Ghedira K..
A Semantic-based Approach for Ontology Module Extraction.
DOI: 10.5220/0004544402220229
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2013), pages 222-229
ISBN: 978-989-8565-81-5

with the entities names inside the ontology.

This paper is organized as follows. Section 2 deals with the previously proposed techniques for ontology modularization. Then in section 3, we present our approach. A case study illustrating the proposed approach is presented in section 4. In section 5, we describe the usefulness of our approach in an application domain namely information retrieval. We conclude in section 6 with a summary of the main points relevant to this study, and we give directions for future work.

## 2 RELATED WORK

Several modularization methods of ontologies have been proposed in the literature. These methods are based on two antagonistic approaches. The first is a composition approach in order to obtain a modular representation. The modularization can be perceived from this perspective as a mechanism for assembling some ontologies (modules) into a coherent network that can be referred to as a single entity, modular ontology. The result is a set of integrated or inter-connected ontologies into a larger and more complex network. The second approach is a decomposition of a large ontology, which contains a large number of concepts and relations into a set of smaller modules, easy to understand and manage.

Decomposition methods proposed in literature, belong mainly to two large families. Partitioning methods are automatic and provide a set of modules that can be disjoint or overlap. Examples include partitioning methods that produce disjoint modules (Cuenca Grau et al., 2007b; Stuckenschmidt and Klein, 2004). Some others, like partition-based methods (MacCartney et al., 2003), allow modules to overlap. As for the extraction methods, they involve the user in the extraction process and provide a single fragment of the ontology. These two categories of methods are generally based either on logical criteria (Cuenca Grau et al., 2007a) or on structural criteria (Ghiraldi et al., 2006; Noy and Musen, 2009; Seidenberg, 2009). In both cases, human intervention is necessary after the modularization process to verify that the module is covering a consistent knowledge area. We believe that this is due to the fact that these methods neglect the semantic aspect in the modularization process.

The methods based on structural criteria target specific ontologies. This is the case, for example, of the method of Seidenberg (Seidenberg, 2009) where the ontology referred, is the GALEN ontology (Rector and Rogers, 1999), which is characterized by the strong presence of hierarchical relationships between concepts. The results of modularization are favourable only in the case of ontologies that have a structure similar to that described at the outset.

The other methods based on criteria of description logic, define the module formally by setting the logical conditions in advance. Portions of ontologies that satisfy these conditions are considered as modules. Although these methods consider a certain level of semantics, the modules are usable only if humans validate the module, by browsing it to estimate the concepts that are relevant to its application. Cuenca Grau et al. propose in (Cuenca Grau et al., 2007b; Cuenca Grau et al., 2005) an algorithm to obtain partitions whose elements are disjoint, starting with a formal definition in order to characterize ontologies that are susceptible to be decomposed safely. Indeed, if the ontology does not have certain formal characteristics defined by the algorithm, it cannot be modularized. This is not ideal because it reduces the number of ontologies ready to modularization. The work of Cuenca Grau et al. is based on the notion of conservative extensions (Ghiraldi et al., 2006; Lutz et al., 2007). This means that essential inferences about entities contained within an element of such a partition should be preserved. Whilst conservative extensions can theoretically be used to define an ontology module, they cannot currently be used in practice as deciding if an OWL-DL module is a conservative extension is undecidable (Doran et al., 2007; Lutz et al., 2007). In (Wandelt and Möller, 2012), the aim is to introduce modularization techniques for ABoxes in order to obtain a set of modules to release the main memory burden of DL reasoning systems for semi-expressive ontologies. They have proposed to transform an ABox to a graph by mapping each individual in the Abox to a node in the graph and then to decompose this graph relying on connectedness-based graph partitioning techniques. The algorithm gave a negative result for *SHOQ* DL (nominals problem, completeness problem). In order to ameliorate results, an intentional-based modularization by splitting role assertions with *ABox-splits* is presented. This method relies on internal paths of role assertions between individuals. The method did not consider the semantic relations expressed by these assertions and the decomposition is completely depending on the graph structure. Furthermore, there are no user requirements considered during the splitting.

Both classes of methods mentioned above reduce the reuse possibility. Indeed, these methods have been dedicated for specific ontologies often characterized by particular structures and properties. In addition, human intervention, for checking the semantic consistency of the concepts that make up the module, is required before using the extracted module by the final application.

In this paper, we propose an approach which take into account the semantic aspect in the modularization process (an application may use the extracted module without the need for human intervention in order to validate it). Moreover, the extraction process does not begin from internal properties of the ontology. When a software agent would extract modules from a set of different ontologies of different knowledge domains, it is not necessary for it to know about the entities inside these ontologies. As a starting point, a set of terms relevant to some domain is entered. The module is produced using a semantic matching between these terms and the ontology concepts. The produced module is intended to be considered as a separate ontology relevant to these terms.



# 3 PROPOSED APPROACH

In general, the modularity of ontologies serves three principal goals:

- The reuse of the fragments (modules) of ontologies in the construction of new ones.
- The interoperability of the distributed systems through the interpretation of the local semantics of ontologies that constitute modules in the global system.
- The extensibility for evolution and maintenance, and the scalability for efficient reasoning by localizing the inference in the module rather than to reason on all the ontology.

In this paper, we propose an approach that serves the purpose of *reuse*. However, reuse here is not intended to assist developers in building new ontologies, as is the case with the other methods of modularization. In fact, we seek primarily to help the user obtaining a relevant ontology module, which captures a set of knowledge from a wider existing ontology. Indeed, it would be interesting to give the user methods and tools that offer an extract from an ontology, which plays the role of ontology in itself. Thus, reusing the module by integrating it directly into an application, saves time to build a dedicated ontology. Note here that the user may be human or machine. It is rather the case of applications that want to use these modules, which interest us the most because we are looking to propose a solution that makes use of modules as ontologies, independently of human intervention.

The modularization approach we propose is part of the decomposition approaches of monolithic ontologies. It is an *extraction method* since it aims to extract a relevant ontology module. The aim of the approach is to provide the user with an ontology module that covers a sub-domain of the domain of the ontology. The method should allow the user to express its needs by entering the concepts which interest him. The result is a fragment composed of concepts and relations that are relevant to the module i.e., which have semantic relationship with the concepts submitted by the user. We consider a semantic relationship between two concepts, as one of the four logic functions as follows:

- *Identity Relation*: it is a semantic relation between two concepts that have the same syntax, the same attributes and operations. Example: Identity (Person, Person).
- *Synonymy Relation*: it is a semantic relation between two concepts that express the same meaning. Example Synonymy (Person, Individual).
- *Classification Is-a Relation*: two concepts where one is expressing a particular case of the other. Example: Is-a (Student, Person).
- *Antonymy Relation*: is used between two concepts that have opposite meanings. Example Antonymy (Registered, Unregistered)

For experimental reasons, we consider only these four semantic relations. These semantic relations exist in WordNet which is a large lexical database for English language. It groups words together based on their meanings and label the semantic relations among words. We exploit these properties to identify the semantic relations between the concepts. For example, in an ontology that describes an e-learning course, the user may be interested in participants in that course. The method should extract a module semantically rich on participants, from the ontology of departure. For this purpose, we verify if one of the semantic relationships described above exists between the keywords entered by the user and the concept of the ontology. The comparison operation is only restricted to named concepts.

We motivate our approach as follows:

- *User Involvement*: In the context of reuse, the user should be satisfied with the result. If not satisfied, he should be able to better communicate his needs to be taken into account in the process of modularization. Our approach involves the user (human or computing system) in the process of modularization, unlike the automatic decomposition approaches (Cuenca Grau et al., 2007b; Stuckenschmidt and Klein, 2004). This allows him to express his needs regarding the result he looks for. He begins by entering the central concepts for the module he wants to achieve. When he gets a result that does not satisfy his needs, he starts the modularization by changing the settings (e.g., concepts to include, concepts to exclude …) to refine the result and get a different module of the previous one.

- *More Semantics*: The extraction methods proposed by Seidenberg, and Noy and Musen involve the user (Noy and Musen, 2009; Seidenberg, 2009). But these methods depend heavily on the structure of the ontology. In addition, the concepts selected by the user are part of the ontology to decompose. In fact, their algorithm follows the links between concepts to determine the portion to be extracted. In our approach, the user may enter concepts that can be internal or external to an ontology. It is a new aspect in the operation of modularization that other methods have not explored. Indeed, these methods work with the concepts that make up the ontology and do not address the case where the user provides concepts that are not elements of this ontology. The essential for us is that the module should capture the meaning of concepts by looking for concepts that are in strong semantic relation with those of the ontology. Thus, the main contribution of our approach is that the module is determined on the basis of the semantic relationship that can exist between internal and external concepts.
- *Low Coupling and High Cohesion*: reuse and extensibility are the goals sought in the operation of an ontology modularization. Nevertheless, achieving these two objectives requires that the modules are loosely coupled and highly cohesive. The coupling means the modules dependency. Loose coupling means a weak relationship between modules allowing flexibility for updating and maintenance. So, each module can be modified by limiting the impact of change to the rest of the ontology. Cohesion measures the dependence of the components of modules. In other words, concepts, relations, and individuals are strongly linked to each other within the same module. So, cohesion denotes the degree of relatedness of elements within the same module (D'Aquin et al., 2009). If we consider two concepts are strongly-related if there is a semantic relationship between them. We can use the semantic relationship, as a mean to identify how strongly-related are the concepts, and consequently if they are parts of the same module. So we can reach high cohesion, in a portion of ontology, based on the notion of semantic relationship.

Before beginning to describe the approach currently being investigated, we propose the following definition of a module: *an extracted ontological module is the relevant part of an ontology to a set of terms which are not necessarily the exact terms used inside the ontology. It is intended to cover a sub-area of knowledge for which a module needs to be extracted.*

This definition implies that ontology module is a single extract and it can be reused as a full independent ontology. The user (human or machine) may be not familiar with the content of the ontology. If he needs to extend the module with new concepts and relations then the module should be viewed as an ontology itself. The quality of the module depends on the relevance of the knowledge captured by the module relative to the user query.

Our approach is based on two basic steps:

- 1st step: Identifying concepts that have a semantic relationship with external terms.

- 2nd step: composition of the module based on the concepts identified in Step 1. All concepts that appear in the definition of the concepts identified are considered part of the module. The module is composed from the union of all retrieved axioms.

The algorithm identifies a module by doing comparisons between a term entered by the user and a concept of the ontology. WordNet is traversed to extract synonyms, antonyms or hyponyms depending on the user choice. In case the term is identical to a concept name, we consider it as part of the module. If there is a semantic relationship between them, the concept of the ontology is moved to the module. In addition, in case the extracted concept has an equivalent definition with another concept in the ontology, all the definition is extracted. So, all the concepts within the module constitute a subset of concepts definitions that are extracted from the original ontology. In case there is not a semantic relationship between the compared concepts, then the ontology concept is not extracted. The algorithm continues so until all concepts in the ontology are compared with the user concept.

At the beginning of the algorithm, the user may choose the concept by entering its name. So, the extraction procedure is automatic but it takes into account the user requirements. In this paper, we present the approach and show its feasibility at a practical level as we show in the next section. The aim of the paper is not to test our method on real well known ontologies. It aims rather at proving the contribution of the modularization based on a semantic matching for some kinds of ontologies (i.e. expressive ontologies not taxonomies). Thus, we present theoretically, in the following paragraph, some evaluation criteria which we can apply on this approach.

As we are only interested in one module, evaluation criteria dedicated to sets of interconnected modules resulting from partitioning techniques – *redundancy*, *connectedness*, and *inter-module distance*–are not relevant in our technique. However, since our method aims to produce a relevant module to a set of terms in order to use the module as an ontology, we can use evaluation criteria for determining the quality of the ontology to evaluate the *quality* of the onotology module. These criteria are mainly the *module cohesion*, the *richness of the representation* and the *domain coverage*.

- *Module* cohesion denotes the degree of relatedness of elements within the module. Cohesion metrics are based on the structure of the ontology: the *number of root classes*, the *number of leaf classes*, the *maximum depth* of the hierarchy.
- *Richness* of the representation denotes the amount of conceptual information retained in the module. The richness of semantic information in a module depends on the richness of the mother ontology. Richness metrics such as - the average number of subclass relations per class- or the -average number of domain relations per class- can be used.
- *Domain* coverage is the criterion which determines how well the module fits the representational requirements of the application that request it. To determine the domain coverage, we need a suitable representation of the domain that should be covered by the module. Comparing a corpus of documents with the module is a technique for determining how well the ontological module represents the content of the documents.

Another evaluation criterion which can be considered is the performance measuring. It is important to consider it, particularly when using a modularization technique for the purpose of an application.

We present in the next section some of the screen shots of our developed system which was tested under an ontology that describes an e-learning course.

# 4 CASE STUDY

We provide an example of extracting a module from an ontology to illustrate our approach. The ontology expressed in description logic corresponds to the following axioms:

a1 correction ≡ page ⊓ ∃associated.exercise

a2 exercise ≡ page ⊓ ∃associated.course

a3 course ≡ page ⊓ ∃caracterized. session

a4 session ≡ ∃caracterize.course ⊓ ∃composed.module ⊓ ∃associated.test

a5 module ≡ ∃associated.Tutor ⊓ ∃associated.registered ⊓ ∃compose. session

a6 test ≡ ∃associated. session ⊓ ∃corrected.tutor ⊓ ∃performed.registred

a7 tutor ≡ ∃associated.module ⊓ ∃correct.test ⊓ teacher

a8 registered ≡ ∃associated.module ⊓ ∃perform.test ⊓ student

a9 person ≡ teacher ⊔ student

Suppose the user wants to extract an ontology module relevant to persons which participate in an e-learning course. He may enter the term “person”, which is the name of an internal concept. He may also enter the terms “coach” or “unregistered”. These words are syntactically different from ontology concepts, but they belong to the same context for the user, that is to say people which participate in an e-learning course.

*Result of the 1st Step:*

If one refers to the semantic relationships defined above, we find that there is an identity relation for the concept person. An antonymy between the concepts registered and unregistered. And a synonymy between coach and tutor.

*Result of the 2nd Step:*

The definitions that we found for these concepts, in the Tbox of the ontology are:

a7 <u>tutor</u> ≡ ∃associated.module ⊓ ∃correct.test ⊓ <u>teacher</u>

a8 registered ≡ ∃associated.module ⊓ ∃perform.test ⊓ <u>student</u>

a9 person ≡ <u>teacher</u> ⊔ <u>student</u>

Note that the concepts that have not a semantic relationship with the original concepts (“teacher” and “student”) chosen by the user appear in the definition of the found concepts. Concepts like “Module” and “Test” are considered as part of the module because they are considered as part of the definition of the concepts founded.

Figure 1. is a screen shot of our developed

system. By clicking on the button Display, all extracted named concepts are listed above and the module is created in a separate owl file (Figure. 2).

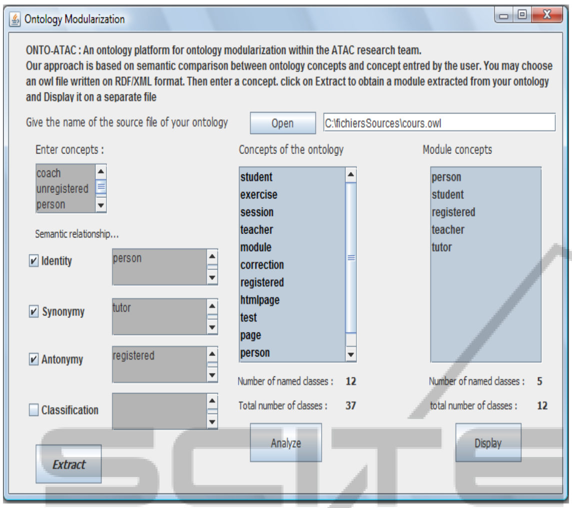


Figure 1: Screen shot of the modularization approach.

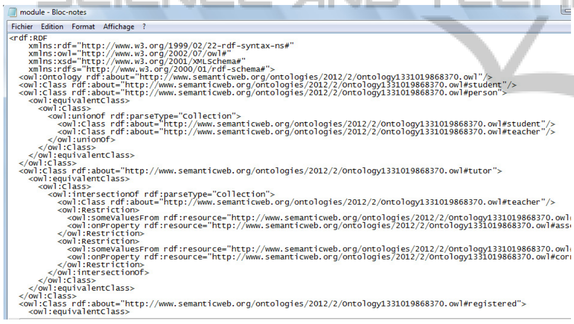

Figure 2: The obtained module in a separate owl file.

## 5 USEFULNESS OF OUR APPROACH

Our approach is based on identifying the semantic relations between terms of concepts. It can have applications in many natural language processing tasks, such as Information Extraction and Information Retrieval. We discuss in this section the usefulness of our approach in the domain of Information Retrieval (IR).

An IR system allows users to look for information in a collection of documents (or other information sources) through queries usually formatted as a set of keywords (Baeza-Rates and Ribeiro-Neto, 1999). There are three main steps for the process of IR: The indexing process, the query processing and the matching between the query

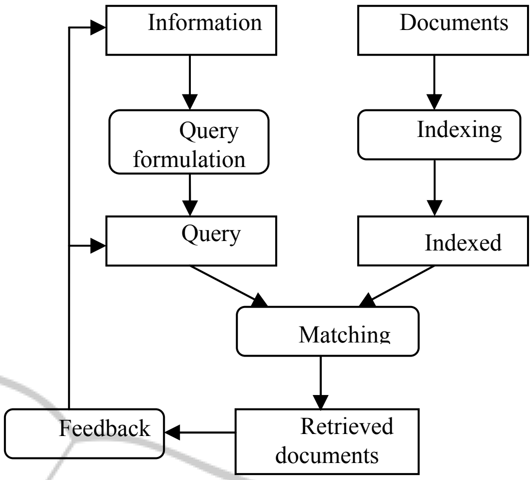


Figure 3: Information retrieval processes.

terms and the documents. These processes are visualized in Figure 3 (Goker and Davies, 2009). The goal of the indexing process is to represent the content of the documents in order to be used in further searches.

There are two types of indexing: bag-of-words indexing and semantic indexing. For the first type, the indexing terms are extracted from the documents content itself. It includes two steps: searching the terms and weighting them. The second type aims to rely on ontologies to represent documents. From this point of view, the descriptors (indexing terms) are chosen directly from the ontology rather than the documents. So, documents are indexed by concepts which reflect their meanings, rather than frequently ambiguous words (Aussenac-Gilles and Mothe, 2004).

Semantic indexing consists of two steps. The first step consists of identifying the ontology concepts or instances in the documents. The second step consists of weighting the concepts for every document according to the conceptual structure which they are derived (Hele and Tanel-Lauri, 2001).

Combining the usability of keyword-based interfaces with the power of semantic technologies is one of the most challenging areas in semantic searching. To use an ontology in an IR system, it needs to choose it first. As much as ontologies in different domains are now accessible, reusing them could be a solution for ontology integration in IR systems. In this case, ontologies are generally chosen only based on the knowledge domain they address (Baziz et al., 2005). Once the ontology chosen, the knowledge it represents can be used when indexing documents. Thus, the choice of the ontologies which will be used for indexing is a primordial step. Our ontology modularization approach would be useful in this context.

In many studies, the choice of the domain ontology which will serve to represent the corpus is dependent of the task domain itself (Vallet et al., 2005). Thus, the reusability of the ontology for another task or another domain is not insured.

In fields, like medicine, ontologies have especially great size, and contain many knowledge domains. A collection of medical documents could be represented by the ontology. We can have a corpus which talks about a specific disease and another corpus which talks about treatment of this disease. As a result of a classic semantic indexing, the two corpuses are indexed by a single ontology. At the end, we obtain many concepts which are shared to represent the two corpuses. This can affect the relevance of the document retrieved later. In this case, our modularization approach would be useful. In fact, in this case, we aim to extract two modules, from this ontology, which are semantically related to the two corpuses. Every module is a representation space of its correspondent corpus. When a query concerning the disease is formulated, only the documents which are indexed semantically by the disease module are retrieved.

## 6 CONCLUSIONS AND FUTURE WORK

In this paper, we proposed a method to extract modules from ontologies based on semantic relations identification. We considered four semantic relations which are: Identity, Synonymy, Classification and Antonymy. We considered that two concepts are relevant for the module if there exists a semantic relation between them. We have used Wordnet to identify the semantic relation between the external concept (the user request) and the internal one (the ontology concept). The result of the extraction is a module composed from these concepts and their definitions.

We show that the use of semantic relations makes the method less dependent to the structure of the ontology to modularize. It is effectively intended to high expressive and more complex ontologies rather than ontology structures based on subsumption relations. The user is involved in the modularization process but he is not supposed knowing the components of the ontology. His needs are expressed as a list of relevant concepts for his purpose. Hence, the method is automatic but takes into account the user requirements. The user here could be a human or an application program. In fact, the main goal of this approach is to allow programs to extract useful modules from available ontologies on the Web. In this way, our goal meets the objective of the semantic Web which is to allow data to be shared, understood and reused across applications.

In future work, we envision to evaluate the usefulness of our approach. For this purpose, we have to determine the possible evaluation criteria, including application-dependent criteria, which can be used to determine the quality of a module. We intend to develop an IR system for medical Web documents using ontology modules to index the documents. The efficiency of the approach would be discussed in the context of experiments that aim to measure the relevance of the retrieved documents.

## REFERENCES

Cuenca Grau, B., Horrocks, I., Kazakov, Y., Sattler, U., 2007a. A Logical Framework for Modularity of Ontologies. In *International Joint Conference on Artificial Intelligence*, pages 298-303.

Cuenca Grau, B., Horrocks, I., Kazakov, Y., Sattler, U., 2007b. Just the Right Amount: Extracting Modules from Ontologies. *In International World Wide Web conference.*

Cuenca Grau, B., Parsia, B., Sirin, E., Kalyanpur, A., 2005. Modularizing owl ontologies*. In Proceedings of the KCAP Workshop on Ontology Management*.

Doran, P., Tamma, V., Lannone, L., 2007. Ontology Module Extraction for Ontology Reuse : An Ontology Engineering Perspective. *In Conference on Information and Knowledge Management CIKM'07,* Lisbona, Portugal.

Ghiraldi, S., Lutz, C., Wolter, F., 2006. Did I damage my ontology? A case for conservative extensions in description logics. In P. Doherty, J. Mylopoulos, and C. Welty, editors, *Proceedings of the Tenth International Conference on Principles of Knowledge Representation and Reasoning*.

Lutz, C., Walther, D., Wolter, F., 2007. Conservative extensions in expressive description logics. In *International Joint Conference on Artificial Intelligence*, pages 453-458.

Noy, N. F., Musen, M. A., 2009. Traversing Ontologies to Extract Views. In Heiner Stuckenschmidt, Christine Parent, Stefano Spaccapietra (Eds.), *Modular Ontologies concepts, Theories and Techniques for Knowledge Modularization,* Springer-Verlag Berlin, Heidelberg, LNCS 5445, pages 245-260.

Rector, A., Rogers, J, 1999. Ontological Issues in using a Description Logic to Represent Medical Concepts:

Experience from GALEN. *In IMIA WG6 Workshop*.

Seidenberg, J., 2009. Web Ontology Segmentation: Extraction, Transformation, Evaluation. In Heiner Stuckenschmidt, Christine Parent, Stefano Spaccapietra (Eds.), *Modular Ontologies concepts, Theories and Techniques for Knowledge Modularization*, Springer-Verlag Berlin, Heidelberg, LNCS 5445, pages 211-243.

Stuckenschmidt, H., Klein, M., 2004. Strutured-based Partitioning of large concept hierarchies. *In Proceedings of the 3rd International Semantic Web Conference.*

Spaccapietra, S., 2005. Report on Modularization of Ontologies. *Knowledge Web Deliverable* 2.1.3.1.

MacCartney, B., McIlraith, S., Amir, E., Uribe, T.E., 2003. Practical Partition-Based Theorem Proving for Large Knowledge Bases. *In Proc. of the International Joint Conference on Artificial Intelligence IJCAI.*

d'Aquin, M., Schlicht, A., Stuckenschmidt, H., Sabou, M., 2009. Criteria for ontology modularization techniques. *In Proceedings of Modular Ontologies*, pages 67-89.

Gruber, T. R., 1993. A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, Vol. 5, 199-220, November 2.

Baeza-Yates, R. and Ribeiro-Neto, B., 1999. *Modern Information Retrieval.* Addison wesley Longman.

Goker, A., Davies, J., 2009. *Information Retrieval Models. In J.Information Retrieval :* Searching in the 21st century. Jhon wiley and sons, Ltd., ISBN-13: 978-0470027622.

Aussenac-Gilles N., Mothe J., 2004. Ontologies as Background Knowledge to Explore Document Collections. *Proceedings of Computer-Assisted Information Retrieval* (Recherche d'Information et ses Applications) - RIAO 2004, 7th International Conference, University of Avignon, France, pages 129-142.

Hele-Mai H., and Tanel-Lauri L., 2001. A survey of Concept-based Information Retrieval Tools on the Web. *Proceedings of the 5th East-European Conference*, ADBI, vol 2, pages 29-41. Vilnius, Lithuania.

Baziz M., Boughanem M., Aussenac-Gilles N., Chrisment C., 2005. Semantic Cores for Representing Documents in IR. *Proceedings of the 20th ACM Symposium on Applied Computing*, ACM Press ISBN :1-58113-964-0, pages 1020-1026.

Vallet D., Fernandez M., Castells P., 2005. An Ontology-Based Information Retrieval Model. *Proceedings of the 2nd European Semantic Web Conference*, pages 445-470.

Wandelt S., Möller R., 2012. Towards ABox Modularization of semi-expressive Description Logics. *Journal of Applied Ontology*, V.7, pages 133-167, DOI 10.3233/AO-2012-0105, IOS Press.