



**HAL**  
open science

## **Modular Ontological Warehouse for Adaptative Information Search**

Nesrine Ben Mustapha, Marie-Aude Aufaure, Hajer Baazaoui, Henda Ben Ghezala

► **To cite this version:**

Nesrine Ben Mustapha, Marie-Aude Aufaure, Hajer Baazaoui, Henda Ben Ghezala. Modular Ontological Warehouse for Adaptative Information Search. 2nd International Conference, MEDI 2012, Oct 2012, Poitiers, France. pp.79-90, <10.1007/978-3-642-33609-6\_9>. <hal-00831648>

**HAL Id: hal-00831648**

**<https://centralesupelec.hal.science/hal-00831648v1>**

Submitted on 7 Jun 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Modular Ontological Warehouse for Adaptive Information Search

Nesrine Ben Mustapha<sup>1,2</sup>, Marie-Aude Aufaure<sup>1</sup>, Hajer Baazaoui Zghal<sup>2</sup>, and Henda BenGhezala<sup>2</sup>

<sup>1</sup> Ecole Centrale Paris, MAS Laboratory, Business Intelligence Team,  
Grande Voie des Vignes 92295 Chatenay-Malabry

{nesrine.ben-mustapha,marie-aude.aufaure}@ecp.fr

<sup>2</sup> Laboratory RIADI, ENSI, La Manouba, Tunisia

hajer.baazaouizghal@riadi.rnu.tn}@riadi.rnu.tn

**Abstract.** With the growth rate of information repositories, most of the current research effort are focusing on improving the accuracy in searching and managing information (especially text data), because of lacking of adaptive knowledge representation to the information content of these systems. Besides, domain knowledge is evolving and consequently, ontologies should be automatically built and extended. Thus, introducing modularity paradigm in ontology engineering is now important to tackle scalability problems. In this paper, we address the problem of representing modular ontologies at an abstract level that can improve the traditional information system with higher efficiency, in the context of previous work aiming at integrating ontology learning in traditional Information Retrieval systems on the web. The contribution consists in organizing ontology elements into semantic three-layered ontology warehouse (topic classification, domain knowledge representation, and module representation). The proposed model has been applied for textual content semantic search and relevance improvement has been observed.

**Key words:** ontology, modularization, knowledge representation, graph-based modeling

## 1 Introduction

With the growth rate of information volume, information access and knowledge management has become challenging traditional Knowledge management System. Domain ontologies are a fundamental form of knowledge representation of a domain of knowledge or discourse [8]. A well constructed ontology can help developing knowledge-based information search and management system, such as search engine, automatic text classification system, content management system, etc, in a more effective way. Then, the performance of these systems depends mainly on two aspects: *Domain vocabulary coverage* of ontologies and *Adaptive ontology representation* for search process over textual data. These aspects have been explored in an anterior work [1] that aims to propose an enhanced

multi-domain information search system based on the integration of incremental ontology learning from users'queries and web snippets. Consequently, additional considerations that have to be taken into account, deal with the scalability of incrementally ontology built and the contextual information requested to use these fragment for Information Retrieval process (IR) (especially query expansion, document filtering and classification). The main objective of this work consists in proposing a modular model of ontologies built from texts, with the aim of using them for multi-domain search on the web. To manage distributed knowledge in a dynamic setting, we need flexible knowledge representation formalism to meet the following requirements for information retrieval: How to define an ontology module according to the semantic of content ? and How to organize multiple networked, distributed and dynamic ontologies over multiple knowledge domains?

In this paper, we propose an abstract syntax of modular knowledge by three-layered ontology warehouse (topic classification, domain knowledge representation, module representation) and definitions of functions that supports navigation between the proposed layers. Since, graphs are a natural way of representing and modeling heterogeneous information in a unified manner (structured, semi-structured and unstructured ones), we choose this structure to design ontology modules. The main advantage of using graphs resides in its dynamic aspect and its capability to represent relations, even multiple ones, between objects. It also facilitates querying and analysis using graph operations. Besides, as there exists a bijection between graphs and DL formalism, then we choose graph-based formalism as an abstract level of modular knowledge representation. The proposed graph-based model is inspired from *attributed Typed graph model* [10] in order to specify a modular ontology by an *attributed typed graph of ontologies*. In the remainder of this paper, we present in section 2, an overview of related works on ontology modularization. Section 3 describes the proposed modular ontological knowledge warehouse for knowledge management systems where formal definitions of the different layers are explained. Section 4 describe the main statements of the graph-based formalism of modular ontologies. Section 5 is concerned with the applicative case study of the proposed model with the aim of performing an adaptative information search. In the next section, the experimentation results are discussed. Finally, we conclude and discuss directions for future research.

## 2 Modular Ontologies

Modularization, in its generic meaning, denotes the possibility to perceive a large knowledge repository (an ontology) as a set of modules, i.e. smaller repositories that, in some way, are parts of and compose the whole knowledge [5]. Several definitions of Ontology Module (OM) exist in the literature, but a consensual definition has not been proposed yet. Therefore, according to some recent works, an OM is seen as a reusable component of ontology, which is self-contained but bears definite relationships to other OM. On one hand, scalability and interpretability of ontologies have generated a significant interest in ontology modularization

from the semantic web community. Recently, there has been growing interest in the topic of modularity in ontology engineering [5]. Ontology modularization is the process of decomposition of a whole large ontology into smaller modules. As a motivating example, we consider one of the case studies of the *NeOn project* [6]- a fishery case study in the Food and Agriculture Organization (FAO) of the United Nations (UN). This case study aims to improve the interpretability of FAO information systems in the fishery domain. However, since the definition for a good software module is already vague [13], there is no well-define agreement on the criteria for decomposing an ontology into modules. Indeed, the definition of an ontology module is still subjective [7] and is intrinsically dependent on the application scenario in which ontology modules are used. Moreover, a number of studies have been carried out on different aspects of modularization (languages for modular ontologies, techniques to extract modules from ontologies, etc.). However, these elements tend to be disconnected from each other and no complete modularization framework have yet been proposed for ontologies, within the application scenario of information retrieval. For these reasons, we intend to provide a generic formalism that would gather under a common framework the different aspects of ontology modularization with the aim of meeting the requirements of a broad range of information search application on the web.

On the other hand, in order to tackle problems of mapping between the ontology vocabulary and the terms contained in textual results, vocabulary and distributional information of terms extracted from documents should be taken into account in ontology representation. In this context, the experimental results of the thesis of Stein L. Tomassen [12] has confirmed the underlying assumption of this work that aims to propose the construction of vectors of context (a set of weighted terms extracted from documents) with the use of domain ontology. The use of these vectors could improve the accuracy of these search systems based on ontologies of 10 %. This proves that the terms extracted from documents indexed can contextualize the search and improve the accuracy of results, despite the presence of a domain ontology. Therefore, information on the distribution of terms in documents are important to be encapsulated in the ontology.

### 3 Modular Ontological Knowledge Warehouse for Knowledge management Systems

Multi-domain ontology modeling is a crucial point due to terminological ambiguity, semantic heterogeneity, knowledge scalability and reuse. In this section, we have defined a modular knowledge representation model, based on three layers (topic level, ontology layer and Module layer). The first layer represents the topic ontology which is a based on a taxonomy of thematic concepts. Each topic is linked to a set of modular domain ontologies (second layer). The second layer represents a set of modular domain ontologies as a network of *Ontology modules* (OM). The third layer is made up with ontology modules definition. An ontology module is an ontology fragment that includes only the concepts and relations that are relevant for a granular knowledge on close concepts. Hence, it is possi-

ble to structure a user query by a conceptual graph made of concepts belonging to a same module. OM can share some concepts and some relations with other Ontology modules.

**A Modular Knowledge Warehouse** ( $\tau$ ) is defined as:

$$\tau = \langle \mathcal{T}, Mo_d, \mathcal{M}, \mathcal{R}, \sigma_o, \sigma_d \rangle \quad (1)$$

Where:

- $\mathcal{T}$  is the topic ontology definition;
- $Mo_d$  is the modular ontology definition;
- $\mathcal{M}$  is the set of module definition composing the modular ontology  $Mo_d$ ;
- $\mathcal{R}$  is a set of Resources' URI;
- $\sigma_o$  is a function associating to each thematic concept of the topic ontology ( $\mathcal{CT}$ ), a modular ontology definition ( $Mo_d$ );
- $\sigma_d$  is a function providing for each instance of module definition ( $Mo_d$ ) with a set of resources ( $\mathcal{R}$ ) (ie. Document URL, Product's URI..).

### 3.1 Topic Ontology Definition

In this ontology, concepts refer to categories of knowledge domain (Art, Business, Health, computer sciences, etc.) linked by mainly subsumption relationships, as proposed by the definition 2.

A **topic ontology** ( $T$ ) is defined as:

$$T = \langle \mathcal{CT}, \mathcal{RT}, \sigma_o \rangle \quad (2)$$

Where :

- $\mathcal{CT}$  is a class definition of topics (domain of knowledge);
- $\mathcal{RT}$  is the signature of an overlapping relationship between thematic concepts  $\mathcal{CT}$ ;
- $\leq^{\mathcal{CT}} : \mathcal{CT} \times \mathcal{CT}$  it is a partial order on thematic concepts  $\mathcal{CT}$  which defines the hierarchy of  $\mathcal{CT}$ ;
- $\sigma_o$  is a function allowing the connection between topic layer and modular ontologies layer by associating to each thematic concept ( $\mathcal{CT}$ ), a modular domain definition ( $MO_d$ ).

This ontology doesn't contain any objects. Only concept and relation definitions are specified. Based on the given definition, an illustration of a topic ontology is described as following:

- $\mathcal{CT} = \{ \text{"Medicine", "Sciences", "Arts", "Business", "E-Commerce", "Home", "Software", "Tourism", "Health", "Internet", "Notebook", "Employment", "Diseases", "News", "Sport"} \}$
- Examples of relations between thematic concepts are  $\sigma_{\mathcal{RT}}(\text{Medicine, Sciences})$  and  $\sigma_{\mathcal{RT}}(\text{Medicine, Health})$ .
- $\sigma_o(\text{Medicine}) = \text{"Medicine.owl"}; \sigma_o(\text{health}) = \text{"health.owl"}; \sigma_o(\text{Sciences}) = \text{"Sciences.owl"}; \text{etc.}$

### 3.2 Modular Domain Ontology: definition

Domain knowledge associated to each thematic concept of the topic ontology (such as human anatomy, human diseases, etc.) is specified by a composition of ontology fragments related to fine-grained areas (etc, head, neck and lung cancers etc.). However, most of existing ontologies follow a monolithic approach, which makes this representation impossible. In our case, a modular domain ontology is composed of modules and inter-module relations. We distinguish two types of relationships: taxonomic and conceptual ones. Because of the general structure of any domain ontology based on a hierarchy of concepts, we propose to define a taxonomic structure of modules by two types of composition: a static and a dynamic composition. The aggregation of ontological modules is called in the literature, "composition". Since the definition of ontological modules is based on the "pivotal concepts", it is possible to deduce the hierarchy of modules from taxonomic relationships between pivotal concepts, basin on an existing domain ontology. Thus, this composition is called "static" because it is based on the predefined hierarchy in the ontology used. Dynamic composition of ontology modules were explored in [11] with the aim of proposing an hierarchical classification of modules based on web-based similarity between pivotal concepts. Inter-module relations are defined using conceptual relations between concepts belonging to overlapped ontology modules. These relationships are characterized by *connection interface*.

A **modular domain ontology** ( $MO_d$ ) is defined as:

$$MO_d = \langle \mathcal{ID}_{Mo}, \mathcal{M}_{i=1..n}, \mathcal{R}_o, \mathcal{H}_M, \mathcal{R}_M \rangle \quad (3)$$

Where:  $\mathcal{ID}_{Mo}$  is the identifier of the modular domain ontology;  $\mathcal{M}_{i=1..n}$  is simply a set of module definitions (described by the formula 4) that are connected by external definitions;  $\mathcal{R}_o$  is a set of roots of ontology modules;  $\mathcal{H}_M$  is a set of taxonomic relationships linking the set of roots  $\mathcal{R}_o$  of Module definition ( $\mathcal{M}$ );  $\mathcal{R}_M$  is an external relation definition between ontology modules  $\mathcal{M}$ .

**Example of modular ontology** An example of modular domain ontology "E-Commerce" is described in Figure 1.

This ontology is described by the following features:

- $\mathcal{ID}_{Mo}$ ="Ontology- E-Commerce";
- The set of modules  $\mathcal{M}_{i=1..4}$ ="person\_Module", "customer\_Module", "product\_Module", "payment\_Modu";
- The set  $\mathcal{R}_o$  of pivotal concepts=pivotalConcept\_ person, pivotalConcept\_ customer, pivotalConcept\_ payment;
- Taxonomic inter-module relations are defined as followings:
  - $Racine\_E - commerce \leq^M person\_Module$ ;
  - $Racine\_E - commerce \leq^M Product\_Module$ ;
  - $Racine\_E - commerce \leq^M payment\_Module$ ;
  - $Module\_Personne \leq^M customer\_Module$ .

Conceptual relations consist of:

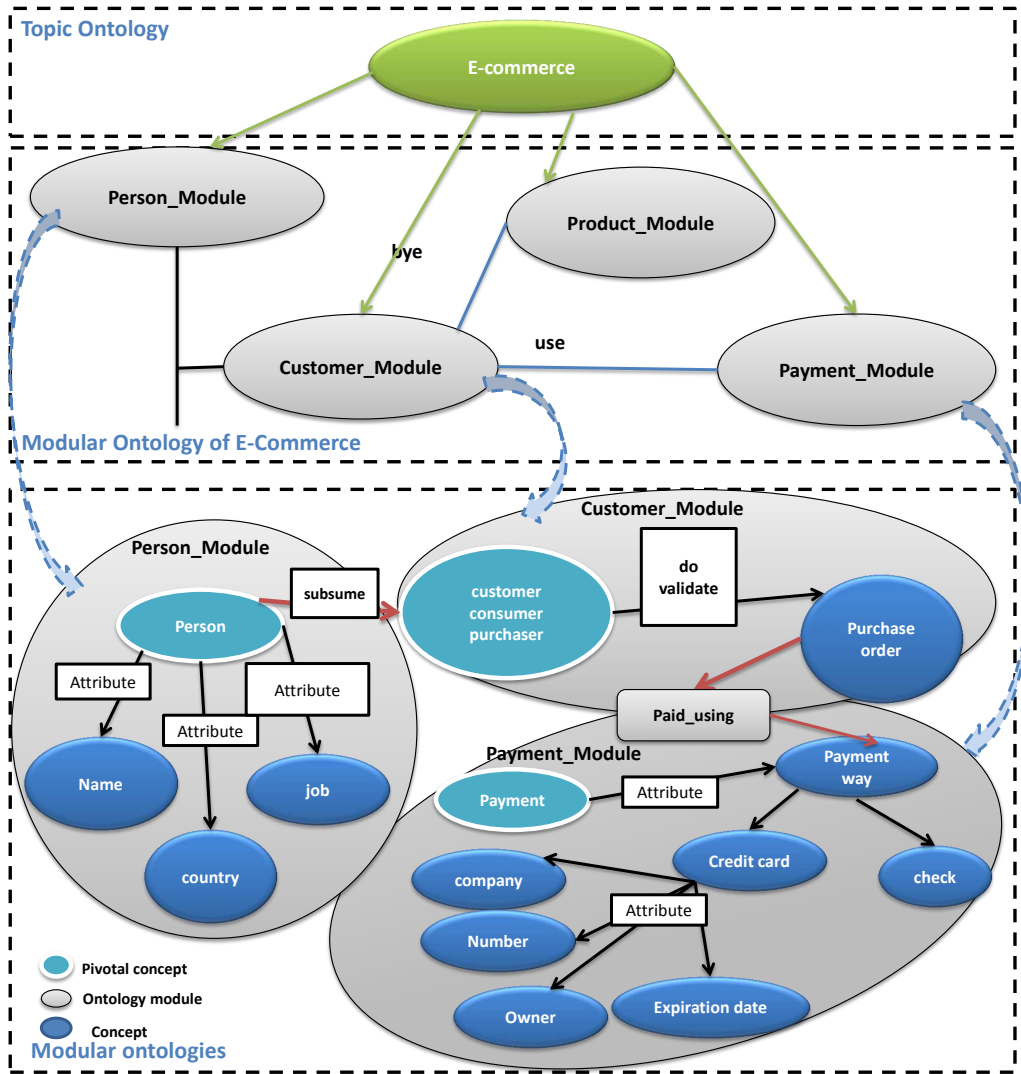


Fig. 1. Example of modular domain ontology " E-Commerce.

- $\mathcal{I}_M(\text{Module\_Person}, \text{Module\_customer}) = \text{Concept\_Person}$
- $\mathcal{I}_M(\text{Module\_customer}, \text{Module\_payment}) = \text{Concept\_type - of - payment.}$

We remind that interface connection between ontology modules is a concept belonging to the linked modules.

### 3.3 Ontology Module definition

An ontology module (OM) is seen as a ontology fragment that has a "meaning", from the viewpoint of applications or users. In the literature, several criteria should be met to comply with the notion of ontological module include: (1) small module size for easy maintenance; (2) independence from the other modules: so that the addition or the removal of a module will not affect many others; and (3) understandability. An OM consists of ontological concepts (referenced by a set of terms) strongly linked by relations (referenced by a set of terms). Each module ontology is characterized by a basic concept, called "pivotal concept" that encapsulates the basic meaning of the granule of knowledge concerned. We define an ontological module by a cluster of related concepts ( $\mathcal{C}$ ) and conceptual relations (CR). The OM encapsulates the ontology of commonsense knowledge elements.

An OM is defined as follows 4.

$$M = \langle \mathcal{ID}_{\mathcal{M}}, \sigma_o^{-1}, \mathcal{C}, \mathcal{R}, \mathcal{O} \rangle \quad (4)$$

where:  $\mathcal{ID}_{\mathcal{M}}$  is the identifier of ontology module;  $\sigma_o^{-1}$  is the inverse function de  $\sigma_o$  that provide a set of thematic concepts ( $\mathcal{CT}$ ) associated to the OM;  $\mathcal{C}$  is a set of identifiers of domain concepts making up the OM;  $\mathcal{R} : \mathcal{C} \times \mathcal{C}$  is a set of relations between the internal concepts of the OM;  $\mathcal{O}$  is a set of objects described by the OMs. Besides, we define the signature of  $(\mathcal{C}, \mathcal{R}, \mathcal{O})$  to be a triple  $(\mathcal{CN}, \mathcal{RN}, \mathcal{ON})$  where  $\mathcal{CN}$  is a set of all names (terms) of concepts defined in  $\mathcal{C}$ ,  $\mathcal{RN}$  is a set of all names (terms) of relations defined in  $\mathcal{R}$ ,  $\mathcal{ON}$  is a set of objects names.

Since the proposed modularization approach is to enhance knowledge search and management, the notion of query is very important to express connection between OMs. This issue is explained in the following subsection.

### 3.4 Internal and external definitions

In contrast with ontology definition, concept and relation inside an OM have local (internal) meaning and distributed (external) meaning. For instance, the meaning of the concept *measure* in the OM related to "body mass Index" in the medical topic is different from its meaning in the chemical topic. Therefore, concept and relation have internal and external definitions.

*Module-based Query* is defined as follows: Let  $\mathcal{VQ}$  be a set of variables disjoint from  $\mathcal{ON}$ . A module-based query  $Q$  over an OM  $\mathcal{M} = (\mathcal{C}, \mathcal{R}, \mathcal{O})$  is an expression of the form of  $q_1, \dots, q_n$ , where  $q_i$  are query terms of the form  $\mathcal{C}(x)$  or  $\mathcal{R}(x, y)$  are such that:

$$x, y \in \mathcal{VQ} \cup \mathcal{ON}, C \in \mathcal{CN} \text{ and } R \in \mathcal{RN}.$$

An **internal concept definition**  $\mathcal{CI}$  is the axiom of the form:  $\mathcal{CN}\mathcal{I} \equiv \mathcal{D}, \{M_i\}_{i:1..n}$  where  $\mathcal{CN}\mathcal{I} \subseteq \mathcal{CN}$ ,  $\mathcal{D}$  is the domain ontology and  $\{M_i\}_{i:1..n}$  is a set of OMs that have in common the concept  $\mathcal{C}$ .

An **external concept definition** is an axiom of the form:  $C \equiv M : Q$  where  $M$  is an OM and  $Q$  is an module-based query over the signature of  $M$  with exactly one free variable.

An **external relation definition** is an axiom of the form:  $R \equiv M : Q$  Where  $M$  is an OM and  $Q$  is an ontology-based query over the signature of  $M$  with exactly two free variable.

As aforementioned, we intend to provide a generic graph-based formalism that would gather under a common framework the different aspects of ontology modularization with the aim of meeting the requirements of a search application on the web.

#### 4 Graph-based Framework for modular ontologies: typed attributed graph of ontologies

At an abstract level, since it exists a bijection between graphs and ontology languages, we choose graph-based formalism as an abstract level of modular ontology definition. Since the OMs are supposed to be extracted from unstructured text, as explained in [4], the discovered concepts and relations are not validated at one step. The ontological representation should take into account change management and automatic update. In our work, we propose to define a general structure of a ontological knowledge base (networked OMs classified by domain) independent of a concrete language. The proposed Ontological knowledge warehouse include a graph-based modeling and functional operators for inter-level navigation. For this reason, we choose to rely on typed attributed graph as it is powerful enough to represent an ontology written in RDF, OWL or DAML+OIL. Besides, typed attributed graph is the model implemented in the ACG library for graph transformation. Details about attributed graph are described in [10].

A **typed attributed graph** representation of the OM  $\mathcal{AG}_M$  is a pair  $(\mathcal{N}_G, \mathcal{E}_G)$ , where  $\mathcal{N}_G$  is a set of attributed nodes and  $\mathcal{E}_G$  is a set of attributed edges. A **typed attributed node**  $\mathcal{N}_G = (T_N, AV_N)$  has a type  $T_N$  and a set of attribute values  $AV_N$  where  $T_N$  is the set of terms referring  $CN$  to a concept  $C$  and  $AV_N$  is the set of score's values assigned to each of the terms belonging to  $CN$ . A **typed attributed edge**  $\mathcal{E}_G = (T_E, \mathcal{RN}, AV_E, O_E, D_E)$  has a type  $T_E$ , a set of attribute values  $AV_E$ , an origin node  $O_E$  and a destination node  $D_E$ , where  $T_E$  denotes the type of a relation (hyponymy, meronymy, possession, etc.) and  $\mathcal{RN}$  is a set of terms referring to the relation defined in  $sig(\mathcal{R})$ . A **typed attribute value**  $AV_E$  is a pair  $(\mathcal{RN}, score)$  associating score's value  $score$  to a term of  $(\mathcal{RN})$ .

In this framework, we propose the following statements:

- Topic ontology is specified by classic typed attributed graph of thematic concepts;
- Modular domain ontology is represented by a typed attributed graph of OMs;
- OM is defined by a typed attributed graph of concepts (ontology definition);

- Each concept is a typed attributed graph of terms.

Furthermore, we propose the model of *typed attributed graph of ontologies* where The nodes and edges are of several types, as follows. The **Term node** (T): is the smallest conceptual unit extracted in the form of a nominal expression. The **Concept node** (C): is a typed attributed graph of *Term* nodes (T) connected by arcs typed as *Terminological relations* (TR) and is a conceptual unit of the *module* (M) graph. The **Module node** (M): is a graph of nodes *Concept* (C) connected with the *conceptual relationships* (CR). The **Ontology node** (GO): is the largest unit defined as a typed attributed graph of *Module* node (M) connected with the *inter-module* interface (I).

### 5 Adaptive Information Search based on Modular Knowledge Warehouse

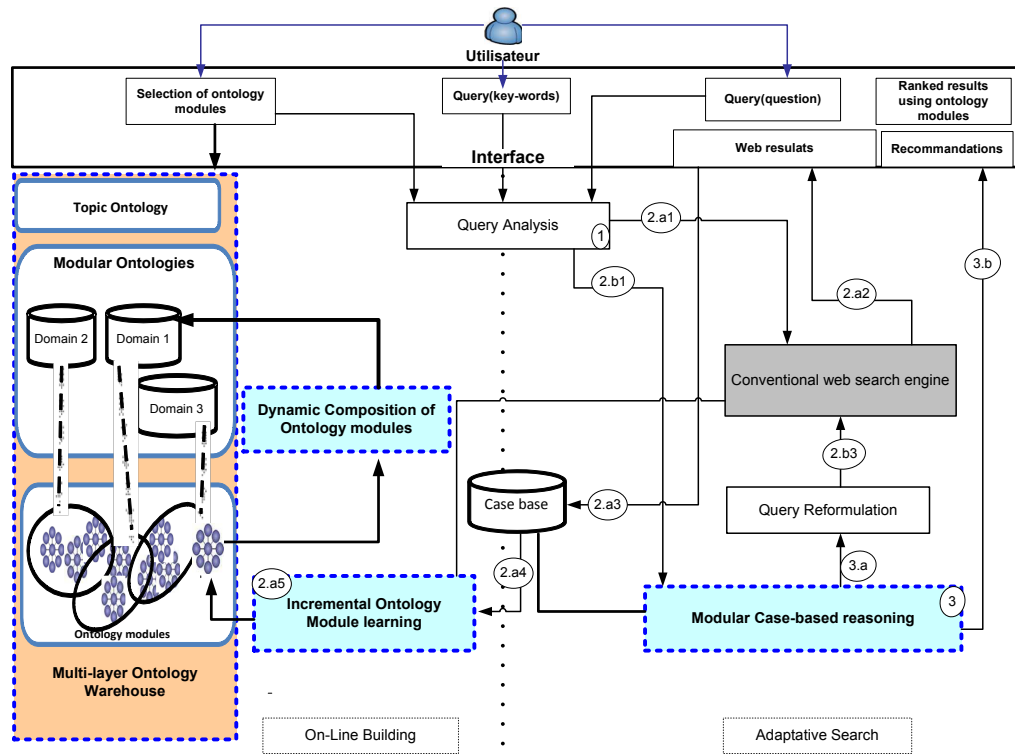


Fig. 2. Adaptive information search based on modular knowledge warehouse

In this view, the user can navigate through the multi-layer warehouse by selecting the corresponding thematic concept and the suitable OM or formulate

a query as a set of keywords or a question in natural language. The query analysis module (1) processes the request according to its type. We distinguish two main cases:

- the absence of an ontology or an OM covering the ontological terms of the request or the response expected (2.a);
- The existence of an OM corresponding to the user query (2.b).

In the first case, the first iteration takes place only by invoking the results delivered by the conventional search engine (ie yahoo, google, etc.) (2.A1 and 2.A2). Then a new case is created in the base case (2.A3) where the query and the results viewed by the user are inserted. This event will trigger the operation of the second component responsible for the learning modules of ontologies from the web (2.A4) [4]. This OM builds the ontology to be displayed in the search interface to allow the user to refine his search. Adding or updating an OM (2.A5) causes the update of the dynamic composition of the modular ontology [11] corresponding modular domain. Therefore, the approach to the *dynamic composition of modules of ontologies* is called to reorganize the concepts being overlap between the different OMs of ontologies and update the hierarchy of OMs based on a semantic similarity measure [11](3.b). In the second case, the component relating to research based on case-based reasoning and modular ontologies can learn a relevance descriptor based on similar cases whose structure depends on the corresponding OM (3). This descriptor will be used to reformulate the query (3.A1) based on semantic proximities (which are calculated based on the number of pages returned ("hit") by the conventional search engine) (3.A2). The results for the reformulated query are provided to the user (3.A1). Simultaneously, the relevance descriptor is used to classify the recommendations relating to OMs semantically close to the current OM and to rank them (3.b).

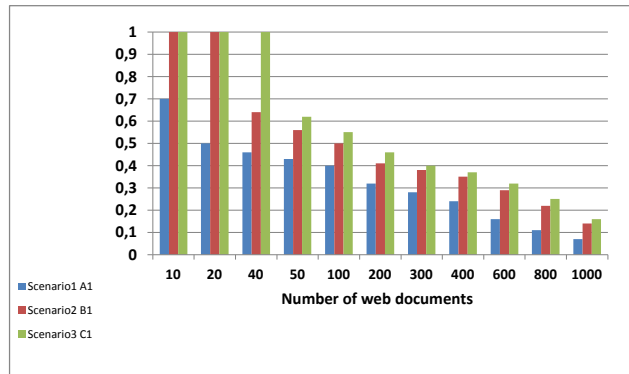
This application is characterized by its adaptivity, for the following reasons. First, it uses ontologies are built incrementally and one based on the needs of users who are looking for information (made by queries). Second, the web search engine will be able to adapt the changement occurring in the areas of knowledge and cooperate users in an indirect way to build the fragments ontologies that represent the common needs of information retrieval **in a modular way**. Finally, the concepts and relationships that make up any OM are weighted by their similarities based on the contextual concurrence with the pivotal concepts of OMs. These weights will be used to reformulate the query with the closest concepts (belonging to the same OM) at a semantic level and at a statistical level (similarity measures based on the number of pages returned by search engines web).

## 6 Experimentation

The main objective of this experiment aims to observe the following aspects of using the multi-layer ontology warehouse: (1) Indexing the case base by OMs for query reformulating based on similar cases, (2) the contribution of document

classification and filtering using the weighted OMs [3] , (3) the contribution of OM enrichment with the search results [4].

Three *scenarios* were conceived. The **scenario A1**: represents a classical search, which involves the search for keywords on Google, as conventional web search. The **second scenario B1** represents the situation where there are similar cases in the database. The research is based on relevance indicator using the weighted OMs to filter the results. The **third scenario C1** represents adaptative information search based on OM enrichment from search results and queries (case base).



**Fig. 3.** L'evaluation de la precision moyenne du module CBRModSearch

In Figure 3, the results show a significant improvement of the relevance of results of returned in third scenario *C1* compared to other scenarios. Indeed, in the second scenario (*B1*), we see a significant increase in the accuracy rate for the first twenty documents. In addition, we also note a significant improvement in results accuracy in the third scenario. This result shows the impact of the reformulation of the query and filtering of documents based on updated OM.

## 7 Conclusion

In this paper, we focus on the necessity of introducing modularity to support the partial reuse of ontologies for semantic annotation in knowledge management systems. We outline some related works of ontology modularization approaches in order to fix the main requirements of the proposed model of modular ontological knowledge. The main propositions described in this paper are: three-layered ontology warehouse (topic classification, domain knowledge representation, module representation). The proposed three-layered ontological warehouse were designed and developed for CBR-based content management where ontology modules are used to annotate cases and compute similarity between cases and recommend

similar content. An improvement of results relevance and content recommendation was observed as described in a previous works. The ongoing work concerns modular and semantic indexing in digital libraries by ontology modules.

## References

1. Ben-Mustapha, N., Baazaoui-Zghal, H., Aufaure, M.-A., and Ben-Ghzala, H.: Combining semantic search and ontology learning for incremental web ontology engineering. In: Sixth International Workshop on Web Information Systems Modeling (WISM 2009), held in conjunction with CAISE'09. (2009) 1036–1049.
2. Ben-Mustapha, N., Baazaoui-Zghal, H., Aufaure, M.-A., and Ben-Ghzala, H.: Semantic search using modular ontology learning and case-based reasoning. In. Proceeding of EDBT/ICDT Workshops. (2010).
3. Ben-Mustapha, N., Baazaoui-Zghal, H., Aufaure, M.-A., and Ben-Ghzala, H.: Enhancing semantic search using case-based modular ontology. In: Proceedings of the 2010 ACM Symposium on Applied Computing, SAC '10. (2010) 1438-1439.
4. Ben-Mustapha, N., Baazaoui-Zghal, H., Aufaure, M.-A., and Ben-Ghzala, H.: Contextual Ontology Module Learning from Web Snippets and Past User Queries. In. Proceeding of KES. Springer LNAI, (2011) 538-547.
5. Stuckenschmidt H., Christine P. and Spaccapietra S.: Modular Ontologies: Concepts, Theories and Techniques for Knowledge Modularization. Springer Lecture Notes in Computer Science (2009)
6. Dzbor, M., Motta, E., Studer, R., Sure, Y., Haase, P., Gmez-Prez, A., Richard V. B. and Waterfeld, W.: NEON LIFECYCLE SUPPORT FOR NETWORKED ONTOLOGIES (OVERVIEW AND OBJECTIVES). In. Integration of Knowledge, Semantics and Digital Media Technology (2005)451–452.
7. M. d'Aquin, A. Schlicht, H. Stuckenschmidt, and M. Sabou.: Ontology modularization for knowledge selection: Experiments and evaluations. In Proceedings of 18th International Conference on Database and Expert Systems Applications, DEXA, volume 4653 of Lecture Notes in Computer Science, Springer pages 874883, Regensburg, Germany, September 3-7, (2007).
8. Gruber, T. R.: Toward principles for the design of ontologies used for knowledge sharing. Technical report, Technical Report KSL-93-04 (1993).
9. Heiner Stuckenschmidt and Michel C. A. Klein: Structure-Based Partitioning of Large Concept Hierarchies. Proceeding of International Semantic Web Conference (ISWC 2004). Springer, Lecture Notes in Computer Science (2004) 289–303.
10. Ehrig, Hartmut and Ehrig, Karsten and Prange, Ulrike and Taentzer, Gabriele: Fundamental Theory for Typed Attributed Graphs and Graph Transformation based on Adhesive HLR Categories. In journal of Fundam. Inf 74 (1). IOS Press(2006)31–61.
11. Elloumi-Chaabene, M., Ben-Mustapha, N., Baazaoui-Zghal, H., Moreno, A., and SAnchez, D. Semantic-based composition of modular ontologies applied to web query reformulation. In Proceedings of the 6th International Conference on Software and Data Technologies ICSOFT (2011). pages 305308. SciTePress.
12. Tomassen, S. L.: Conceptual Ontology Enrichment for Web Information Retrieval. Phd, Norwegian University of Science and Technology NTNU (2011).
13. D.L. Parnas.: On the Criteria To Be Used in Decomposing Systems into Modules. Communications of the ACM, 15(12), December (1972).