International Journal of Information, Control and Computer Sciences
ISSN: 2517-9942
Vol:4, No:6, 2010

# Toward a use of ontology to reinforcing semantic classification of message based on LSA

S. Lgarch, M. Khalidi Idrissi and S. Bennani

***Abstract*—** For best collaboration, Asynchronous tools and particularly the discussion forums are the most used thanks to their flexibility in terms of time. To convey only the messages that belong to a theme of interest of the tutor in order to help him during his tutoring work, use of a tool for classification of these messages is indispensable. For this we have proposed a semantics classification tool of messages of a discussion forum that is based on LSA (Latent Semantic Analysis), which includes a thesaurus to organize the vocabulary. Benefits offered by formal ontology can overcome the insufficiencies that a thesaurus generates during its use and encourage us then to use it in our semantic classifier.
In this work we propose the use of some functionalities that a OWL ontology proposes. We then explain how functionalities like "ObjectProperty", "SubClassOf" and "Datatype" property make our classification more intelligent by way of integrating new terms. New terms found are generated based on the first terms introduced by tutor and semantic relations described by OWL formalism.

***Keywords*—**Classification of messages, Collaborative communication tools, Discussion forum, e-Learning, formal description, Latente Semantic Analysis, Ontology, OWL, semantic relations, Semantic Web, Thesaurus, tutoring.

## I. Introduction

To collaborate in a tutoring system in e-Learning, the use of collaborative tools is essential. The tutor and learners can communicate via synchronous or asynchronous tools in particular discussion forum thanks to guarantee freedom in terms of time, because they don't require the presence of all players in the same slots time for communication.

However, sometimes this type of collaboration tools is not easy to handle, when the volume of messages accumulated over time in a progressive manner. This makes the exploitation of communication space very complex. Hence the need for tools of classification and organization to facilitate searching and to help tutor to access to information in a simpler manner.

To help a user who can be a tutor or an instructor to find a message posted in a discussion forum, most classification methods provides a search based on keyword. The research results are dependent and proportional to the appropriateness of terms used for search.

An approach to manage this mass of messages, by a classification of messages according to their semantic context was presented in [1]. This classification is based on the method LSA (Latent Semantic Analysis). The construction of a thesaurus that will bring to the messages posted by learners, a semantic context was also proposed. The results thus found by using a thesaurus seem satisfactory [1]. However it is necessary to signalize some insufficiencies in using the thesaurus and that we saw from results found.

The purpose of this paper is the use of ontology to organize the terms of our vocabulary in a very clear way. That organization is based on the functionalities that ontology disposes, in particular the property of its formal aspect. Once our ontology is built, it will be used in the research phase of semantic similarity between terms entered by the user and those that ontology organizes, and so the classification of messages will be richer semantically.

The following plan will be adopt. In Section 2 we mention the importance of collaboration tools for tutoring especially those asynchronous as the discussion forum, while presenting the problem that this type of tool generates. We then describe the essential elements on which is based the semantic classification tool presented in [1]. Section 4 is devoted to presenting the approach adopted to allow the classification semantics of messages and also the improvement brought to this approach while interpreting the results. The insufficiencies identified in the use of the thesaurus will be the subject of the fifth section. In Section 5, we also present the advantages that ontology possesses for overcome these insufficiencies. In Section 6 we introduce the notion of ontology. The use of some functionalities that OWL provides to make the classifier more intelligent will be explained in Section. At the end we give a conclusion and prospects for our next works.

## II. Collaboration between the tutor and the learner by way of asynchronous tool

The success of any work performed by several actors who have to work together to achieve a common goal, depends on collaboration tools available to them. When the work to achieve successfully is a work of distance learning, success becomes a challenge for all intervenors in this work. There where collaborative learning is organized by interactions both

S. Lgarch is a Phd student at Mohammadia Engineering School, Mohamed V– Agdal University, Rabat, Morocco (e-mail: saadiagm@gmail.com)
M. Khalidi Idrissi is a professor at Mohammadia Engineering School, Mohamed V– Agdal University, Rabat, Morocco (e-mail: khalidi@emi.ac.ma)
S. Bennani is professor at Mohammadia Engineering School, Mohamed V– Agdal University, Rabat, Morocco (e-mail: sbennani@emi.ac.ma)

synchronous and asynchronous, between learners and their tutor has shown its advantages in the success of online formation [2] [3].

Given the importance of the side tutoring for any device of distance learning and the role played by the tutor to overcome the problem of isolation that the learner may feel and that presents a real obstacle in the continuity of his formation [4] the need to use communication tools is essential.

The asynchronous communication tools, particularly discussion forums allow the exchange of information in flexible way. But in return they generate a large mass of messages. We thus see that the volume of messages exchanged generates noise, proportional to the number of interveners. This makes the exploitation of this mass a heavy and impractical. The undesirable mixture of messages from different contexts and different objectives generates a block and slowness in reply's time. A member of a working group that is remote, Requires functionalities to be included in the asynchronous communication tools to facilitate to him the task of researching the desired information in a very fast way and depending on the intended context [5].

A tool for semantic classification of messages of a discussion forum was proposed in our work presented in [1].

## III. Basic Element of our Classification Tool

The classification tool introduced in [1] is based on LSA (Latent Semantic Analysis) with a reinforcement of the classification by integrating a thesaurus.

Based on the singular value decomposition (SVD), the LSA method allows to find similarities between the documents (texts, sentences, words) [6][1].

In order to have relevant results we have proposed to widen the scope of research while respecting the context requested. The use of the technologies proposed by the Semantic Web in particular those that enable the organization of vocabularies in a semantic way, was necessary. For this, we chose the thesaurus.

### *A. Semantic web*

The term Semantic Web attributed to Tim Berners-Lee [9] denotes a set of technologies to make the content of resources on the World Wide Web accessible and usable by software agents and programs, through a system of formal metadata, including using the family of languages developed by W3C.

The Semantic Web does not call into question the classic web, because it is based on it, especially a means of publication and consultation documents. The automatic processing of documents via the semantic web is done by adding formalized information (markers) that describe their content and their functionalities instead of texts written in naturals languages (French, Spanish, Chinese, etc..) [6]. Moreover, for the manipulation of semantic markers, we need semantic resources that help to define a vocabulary for such markers and also allow concepts sharing and interoperability. Among these resources we find the taxonomies, semantic networks, thesaurus and ontologies [1].

### *B. Thesaurus*

The international standard ISO 2788 (1986) defined the thesaurus as the « vocabulary of a controlled indexing language formally organized in order to explicit the relations priori between notions (eg relation generic / specific) ». According to the same standard, an indexing language is a « set of controlled terms and selected from a natural language and used to represent in condensed form, the contents of documents ».

The thesaurus was designed in the late 1950s. Its first function was to overcome the disadvantages of natural language: by grouping different meanings in the same form meaningful and dispersion of information in terms more or less similar semantically. The thesaurus is as an instrument of control and structuring of the vocabulary; it contributes to the consistency of indexing and facilitates information retrieval [7].

The terms in a thesaurus are conceptually organized and interconnected by semantic relations. These relations are of three types: hierarchical, equivalence and association [1].

The possibility that the thesaurus gives in terms of semantic classification of terms of a given vocabulary, we have encouraged on one hand to integrate it as an essential component in the classification presented in [1]. On the other hand, the simplicity of relations and of terms that the thesaurus presents has facilitated the implementation of the classifier and to see the first results when a semantic resource of organization of words is integrated

## IV. Classification of Messages of a Discussion Forum Based on the LSA

To help a user find a message posted to a discussion forum, most methods of classification provides a research based on keywords. The research results obtained are dependent and proportional to the relevance of keywords chosen by the user.

A tool for classifying the messages of a discussion forum that is based on a semantic approach was presented in [1]. This approach allows managing the mass of messages accumulated with applying a classification according to their semantic context. The classification made is based primarily on the LSA method. In order to increase the performance of the method chosen by extending the terms used in the construction of Table lexical (words / documents) and thus improve the classification, we thought to organize these terms with other terms in a hierarchical manner using a thesaurus.

Our implementation was done in three stages. In the first one, we only implemented the LSA. The object of the second stage was the implementation of our basic semantics approach and that we improved in the last stage.

### *A. Using the LSA only*

Having defined a context of classification of messages using a set of terms (keywords), the first test done on the classification is based solely on the LSA, demonstrates restrictions on the results and which are due to the statistical

nature of LSA method. The results thus found from this classifier ignore messages from the same desired semantic context, if they don't contain any keyword defined on starting.

### B. Integration of the thesaurus

The integration of the thesaurus as a semantics resource has been the subject of two approaches. The first approach consists to include in addition to keywords specified by the user, the specific terms that are associated to them through the thesaurus, avoiding repetitions [1]:

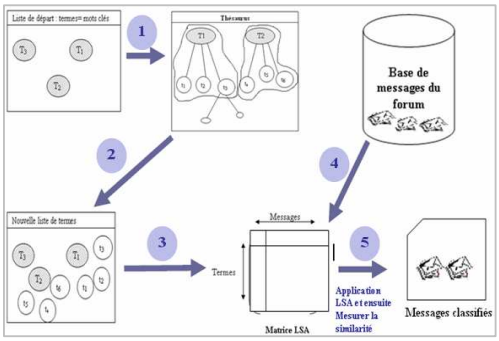


Fig. 1 – General architecture of the system

This approach demonstrates that the results generated are more interesting in terms of semantics as those generated by the LSA method only, because messages with semantics near to that desired are generated, without these messages contain the specified keywords. But some messages of different semantics are also returned, since they contain terms that are related to a few key words only and not to all of these keywords [1].

To overcome the problem of side messages, an improvement to semantic approach of classification is made. In this case and to build the lexical table, we include in addition to the keywords specified by the user, specific terms defined by the thesaurus, common to those [1]:

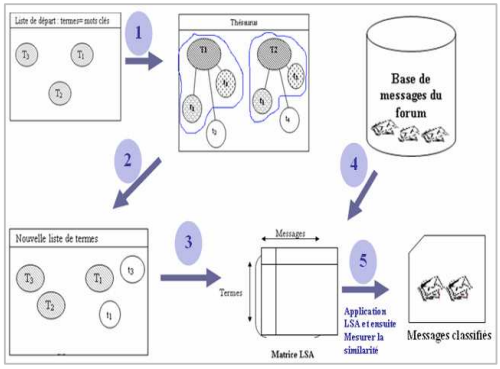


Fig. 2 – lexical Table including only the common terms

The Improvement made to our basic approach leads to more relevant results than those generated from the first approach. The messages returned are only in the same desired context.

The improved semantic approach allows classifying messages according to a set of terms that belong to the desired themes, based on semantic relations that exist between these terms. The terms used so to enable this classification, are ranked according semantic relations using a thesaurus. The latter is constructed from a corpus of messages of different topics. The application of this approach on a corpus of messages posted through a forum discussion, showed results relevant and rich in semantics, which approves the use of thesaurus prior to the LSA.

### C. Interpretation of results

To compare the three implementations, we calculate statistics on all the search results. So for a corpus of 115 messages, on which we applied a classification based on the theme: "routing" and "protocol", we learned the following results [6]:

TABLE I
RATE OF MESSAGES RETURNED WHICH RELATE TO THE THEME PROTOCOL OF ROUTING

| | Number of messages returned | number of messages of the corpus similar to the chosen themes | number of messages returned and similar to themes chosen | Rate of messages returned and similar to the chosen themes |
|---|---|---|---|---|
| LSA only | 27 | 16 | 9 | 56,25 % |
| Basic approach | 85 | 16 | 16 | ≃100 % |
| Improved approach | 33 | 16 | 16 | ≃100 % |

The TABLE I show firstly that the proposed semantic approach, gives more interesting results in terms of semantics as those given by the LSA method only and that, by achieving a rate close to 100% while the rate achieved by the LSA does not exceed 56.25% [6].

On the other hand, TABLE II shows that our second approach is more attractive to overcome the problem of spam and this, by achieving a rate of 51.52% of spam compared to the results returned by the basic approach and the LSA which carry rates of respectively 81.18% and 66.67% [6].

TABLE II
RATES OF SPAM RETURNED

| | Number of messages that are not similar to the chosen themes | Number of messages returned | Rate of messages returned that are not similar to the chosen themes by the total of messages returned |
|---|---|---|---|
| LSA only | 18 | 27 | 66,67 % |
| Basic approach | 69 | 85 | 81,18 % |
| Improved approach | 17 | 33 | 51,52 % |

The results thus found by using a thesaurus seem satisfactory. However it is necessary to highlight some insufficiencies in the use of thesaurus.

## V. Insufficiencies of thesaurus

The thesaurus has been created to assist archivists in their task of indexing and queries formulation [14]. It's characterized by a degree of semantic precision given for the presentation of knowledge that limits its use for automatic indexing. This is explained partly because a terminology dictionary, incarnates a representation of a domain (a lexicalization of a conceptualization), which is not as complete as the formal semantics provided by the conceptual representation, and its modest structure, is therefore unsuitable for advanced semantic applications. On the other hand, and in particular, relations linking terms (controlled vocabulary to represent concepts) in a thesaurus (BT, NT, RT) are generally not sufficient for a profound analysis of the semantics of indexed documents [17].

The thesaurus also lacks a conceptual level of abstraction. These are collections of terms that are organized under a single hierarchy or multiple hierarchies but with basic relations between terms. The distinction between a concept and its lexicalization is not clearly established. The thesaurus does not reflect how the world can be understood in terms of meaning. In addition, coverage semantic thesaurus is limited. The concepts are generally not differentiated from their abstract type (such as substances, processes). The relations between terms are vague and ambiguous. The relation "is related to" is often difficult to exploit because it connects the terms by implying different types of semantic relations. It is often difficult to determine the properties of relations "more specific», «more generic» which can combine the relations «is an instance of» or «is part of». The thesaurus also lack consistency and may contain conflicting information [14].

The gains made by reuse, are many. It was perceived for a long time as a means to improve quality and reduce costs and delays in production. Yet like in other areas, reuse in e-learning has become a discipline and focus of research in its own right [13]. In this context, we are interesting to the reuse of knowledge bases, something that a thesaurus can not satisfy.

An investigation on the side of the ontology is then conducted. Ontology allows reuse by creating and maintaining reusable knowledge. It allows also the assembly of knowledge bases from reusable modules. The sharing of knowledge and communication is also possible with ontologies since they provide interoperability between systems and enable the exchange of knowledge between these systems [8].

Ontology can thus overcome the insufficiencies of the thesaurus through the opportunity to represent the knowledge of a domain by identifying and modelling concepts and conceptual relations. The ontology can also formalize the conceptualization and corresponding vocabulary, this formalization which also targets to remove any ambiguity [16].

All these qualities that ontology possesses render its degree of semantic precision for the presentation of knowledge higher. An adaptation of our classifier to ontology instead of a thesaurus is then proposed.

## VI. Ontologies

Ontology is an explicit specification of a conceptualization of a domain, formed by concepts and relations that allow humans and machines have everything they need to understand and reason about an area of interest or a portion of the universe [11]. On one hand, ontologies allow to describe the knowledge of a specific area, and on the other hand to represent complex relations between concepts, axioms and rules [12]. Ontologies have become a central component in many applications, and they are called to play a key role in building the future "Semantic Web" [10].

A thesaurus or even a taxonomy are forms of ontology whose grammar has not been formalized. When we establish a category and a hierarchy of this categorization, we establish dependencies between these terms. These hierarchies are meaningful outside the vocabulary itself. For example, when we say «this term is a subcategory of that other term», we come giving sense of this relation, we draw a "arrow" between the two by qualifying the arrow and asserting what kind of relation that meant. Ontology corresponds therefore to a controlled and organized vocabulary, and to explicit formalization of relations established between the different vocabulary terms. To realize this formalization, we can use a particular language. Among the languages used to describe the relations between various terms of vocabulary, there are RDF(S) and OWL [15]. All the benefits listed above and relating to ontologies encouraged us to propose a future work using ontology instead of a thesaurus for controlling our vocabulary.

## VII. Adaptation of our semantic approach with a ontology

The adaptation of the semantic approach will be at the level of the search of the new terms organized by ontology and that we chose for replacing the thesaurus. In this paper, the contribution of semantic relations that can exist in an ontological organization in the process of classification is especially focussed.

### *A. Building a ontology core*

The construction of ontology test is our first step. To achieve this operation, we are working on a whole corpus of messages of different thematic (routage, protocole, IGP, dynamique, web, etc...) and that are posted via the platform of distance learning moodle [6]. To extract terms that contain information about the messages of the corpus, we followed the same procedure as explained in [1]. The terms found are organized according to superclass and subclass hierarchy ("routage" is a superclass of "protocole_routage",

"prorocole_securite" and "protocole_application". The term "protocole_routage" is a subclass of "routage") (Fig. 3). Then, we assign to each final class (a final class has no subclass), its own instances. ("protocole_routage" has "IGP", "BGP", "OSPF", and "RIP" as instances). After the operation of hierarchical organization, we proceed to define the semantic relations that may exist between the concepts of ontology, and that we can create according to our intentions, unlike the thesaurus that does not allows the representation of such relations. For example, we have define tree semantic relations "utilise_connecteur", "utilise_metrique" and "utilise_protocole" between "dynamique" and "connecteur_reseau", "metrique", "protocole_routage" respectively.

To create the core ontology, we chose the ontology editor "protégé" in the version 3.4.1. Protégé is an ontology editor that allows the development of OWL ontologies and other forms. Its interface is very intuitive and the software is fairly mature [18].

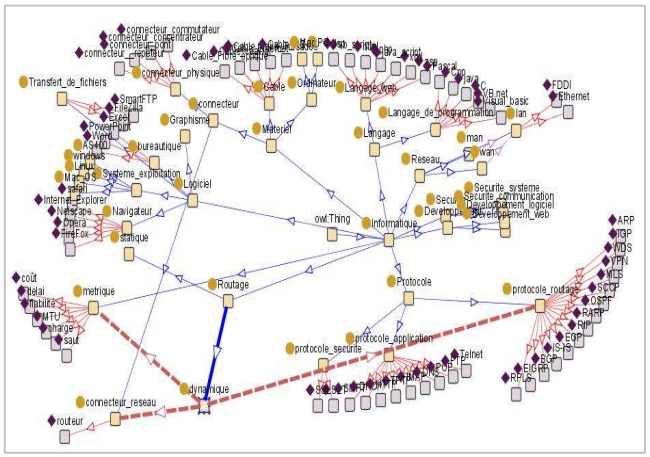


Fig. 3 – Organization vocabulary following the ontology structure

After building the ontology core using "protégé", it will then be exported to OWL formalism, something that the ontology structure allows unlike thesaurus. The choice of OWL language came after showing that it has more advantages comparing to RDF(S) [6]. The OWL file generated by "protégé", expresses clearly and unambiguously the semantic relations that excite between the concepts of the ontology core. In figure (Fig. 4), the OWL code explicit examples of semantic relations between the "dynamique" concept and tree others concepts that we described in (section 7.1). We can also observe that the relation "has Subclass" is automatically generated by OWL code when we define a subclass of a class.

```
    <owl:Class>
      <owl:unionOf rdf:parseType="Collection">
        <owl:Class rdf:about="#Ordinateur"/>
        <owl:Class rdf:about="#Cable"/>
        <owl:Class rdf:about="#connecteur_physique"/>
      </owl:unionOf>
    </owl:Class>
  </rdfs:range>
 </owl:ObjectProperty>
<owl:Class rdf:ID="dynamique">
  <rdfs:subClassOf>
    <owl:Class rdf:ID="Routage"/>
  </rdfs:subClassOf>
 </owl:Class>
 <owl:ObjectProperty rdf:ID="utilise_connecteur">
  <rdfs:range rdf:resource="#connecteur_reseau"/>
  <rdfs:domain rdf:resource="#dynamique"/>
 </owl:ObjectProperty>
 <owl:ObjectProperty rdf:about="#utilse_protocole">
  <owl:inverseOf rdf:resource="#est_utilise"/>
  <rdfs:range rdf:resource="#protocole_routage"/>
  <rdfs:domain rdf:resource="#dynamique"/>
 </owl:ObjectProperty>
<owl:ObjectProperty rdf:ID="utilise_metrique">
  <rdfs:domain rdf:resource="#dynamique"/>
  <rdfs:range rdf:resource="#metrique"/>
 </owl:ObjectProperty>
 <owl:ObjectProperty rdf:ID="necessite_outils">
  <rdfs:comment
```

Fig. 4 – Example of semantic relations formalised by OWL language

The clarity of the formal representation of the vocabulary that the ontology suggests in particular the OWL, allow a query more exact for this semantic resource.

### *B. Toward the use of a set of functionalities proposed by the formalism OWL*

The formal ontology is very rich in semantic relations that may exist between concepts. The important advantage of this formal presentation is the way in whish knowledge is presented with à wide clarity and more precision, hence the absence of any ambiguity of the treated vocabulary

The integration of ontology in the semantic classifier will play an important role in a targeted research of new terms related with semantic relations more exact than those proposed by a thesaurus. These terms are so well represented using the OWL formalism chosen. The general architecture of the classification tool remains the same the only changes will be made at the stage of semantic similarity search using the OWL ontology.

For a better organization of our vocabulary's concepts through ontology, we made use of compound terms. In this case and in order to preserve the semantic that gives each sub-term, we propose to integrate all sub-terms in the classification phase (exactly in the construction of the lexical table LSA) with avoiding repetitions in the set of terms integrated.

The core ontology built, and especially its OWL formal aspect gives a set of functionalities to get easy the manipulation of diverse relations that may exist between two terms. In this work we propose the use of "ObjectProperty", the "SubClassOf", and "Datatype" property.

"ObjectProperty"is used to describe semantic relations between two terms, for example relation "utilise_protocole" is defined between "protocole_routage" and "dynamique" (figure 4). In this case the introduction of "dynamique" as key word

by user will make the classification process more intelligent thanks to the "ObjectProperty" functionality and its use to generate "protocole_routage" as new term.

With OWL ontology, it's also possible to make a filtered classification of messages by the way of the incorporation only the terms that cover the query of user. The accuracy of relations provided by a formal ontology and defined between two concepts, ignored the implication of terms that are semantically far from the theme initially introduced by the user. For example the concept "protocole_routage" is a subclass of "protocole" as "protocole_securite" and "protocole_application", but these last two concepts are not called in the classification stage because the semantic relation " utilise_protocole" is exactly defined between "dynamique" and "protocole_routage and thus the classifier allows filtering under the following theme desired by user.

The relation between a class and its subclass is automatically created. For example, we find relation between "routage" and its subclasse "dynamique". In this case, a simple introduction of "routage" as key word by user, generate the call for "dynamique" which is linked to "routage" through the tag "SubClassOf" of OWL language. This type of relation between classes can increase the quality of the classification tool that will be more intelligent. The increasing of intelligence of the classification is possible through integration of new terms in the classification phase. In this case, instead of using solely on the concept, we also appeal to all of these sub-concepts and also their instances. The introduction of "routage" as a key word by user will be accompanied by the integration of "dynamique" and "statique" in the matrix LSA without forgetting the appeal of all instances of the two subclasses.

"Datatype" property can be a solution to overcome some difficulties of the manipulation of the natural language in whish messages are wrote. Among those difficulties, we cite the case of messages that are wrote in French and that contains some words from English language. For this and for not losing semantic information that these English words give to the message, we propose the use of a "Datatype" property assigned to French term and that will takes its English translation as value. This proposition will give more intelligence to the classification thanks to the widening of semantic area which is presented by the French terms and their translation.

To have a valid OWL ontology in its Full version, we used non-accented terms for defining concepts, instances and relations. But when messages contain some words that are accented and that give information to messages, their neglect will generate a remarkable loss of semantic, hence the need to integrate non-accented words in the classification phase. The use of a datatype property is our proposition. For this, we assign for each non-accented word a "Datatype" property and which the value will be the accented word.

## VIII. Conclusion and prospects

We have presented in this paper a set of functionalities that a formal OWL ontology proposes.

The functionalities explained, can widen the field of research by integrating new terms emerged from the ontology thanks to the semantic relations of type ObjectProperty (eg "utilise_connecteur" and "utilise_metrique"), semantic relations of type SubClassOf (between "routage" and "statique") and datatype (eg "traduction").

The tests that we proposed in this work, and the implementation of others functionalities that ontology provides, like the opportunity to be interviewed by a query language for example SPARQL, will be the subject of a future work. The future implementation of the interrogation process of the OWL ontology, we will use the Jena Framework. Jena is dedicated to building semantic web applications. It allows the manipulation of ontologies by providing Java APIs [19].

Reuse is also a strong point of ontology, and in this prospect the core ontology already created will be fed permanently with new terms to allow its reuse in other projects.

The proposition to adopt a Service-Oriented Architecture which is based primarily on the potential of a combination of XML, Web, specifications of SOAP and WSDL, which were designed to promote interoperability and extensibility, will be also a subject of our future work.

## References

[1] S. Lgarch, D. Bouzidi, S. Bennani « Une approche sémantique de classification de messages d'un forum de discussion basée sur la méthode LSA » 1er Workshop sur les Nouvelles Générations de Réseaux : La Mobilité, WNGN'08, FST de Fès, 2008.

[2] F. Henri, K. Lundgren-Cayrol " Apprentissage collaboratif à distance, téléconférence et télédiscussion", rapport interne n°3.Montréal, Canada; Centre de recherche LICEF, 1996.

[3] M. Walckiers, T. De Praetere "L'apprentissage collaboratif en ligne, huit avantages qui en font un must ", distances et savoirs 2004 (Volume 2), ISSN 1765-0887, page 53 à 75,2004.

[4] B. Charlier, A. Daele, F. Docq, M. Lebrun, S. Lusalusa, R. Peeters, N. Deschryver « tuteurs en ligne : quels rôles quelle formations », in Actes du Symposium international du C.N.E.D, Poitiers, 1999

[5] D. Bouzidi « Collaboration dans les cours hypermédia du système de télé- enseignement SMART-Learning », thèse de doctorat, Université Mohammed V- Agdal -Ecole Mohammadia d'ingénieurs, 2004.

[6] S. Lgarch, M. Khaldi Idrissi et S. Bennani "Adaptation of a classification semantic approach of messages of a discussion forum based on the LSA method to ontology". IEEE SIIE '10:In the proceedings of the 3rd International Conference on Information Systems and Economic Intelligence , IEEE CS Press, Sousse, Tunisia, 2010.

[7] L. Saadani, S. Bertrand-Gastaldy, " Cartes conceptuelles et thésaurus : essai de comparaison entre deux modèles de représentation issus de différentes traditions disciplinaires", Canadian Association for Information Sciences, Procedings of the 28th Annual Conference, 2000.

[8] J. Charlet « Ontologie », STIM/DSI/AP–HP & Université. Paris 6 & groupe « Terminologie et Intelligence Artificielle » Séminaire ISDN/Web sémantique, 24 mai 2002

[9] T. Berners-Lee, J. Hendler, O. Lassila, "The semantic web", A new form of web content that is meaningful to computers will unleash a revolution of new possibilities. Scientific American, pages 35–43, 2001

[10] M. Boutet, A. Canto, E. Roux « Plateforme d'étude et de comparaison de distances conceptuelles », Rapport de Master, Ecole Supérieure En Sciences Informatiques, Université Nice Sophia-Antipolis, 2005

[11] A. Moujane, D. Chiadmi et L. Benhlima. "Semantic Mediation: An Ontology-Based Architecture Using Metadata". CSIT2006, Amman, Joranie, Avril 2006.

[12] H. Zargayouna " Indexation sémantique de documents XML " Thèse de doctorat, Université Paris-Sud, 2005.
[13] H.Kara Terki, A. Chikh « Patterns d'analyse pour l'ingénierie des besoins en e-learning », Rapport de Magiter, Université de Tlemcen, Faculté des sciences de l'ingénieur, 2007
[14] Hernandez N. (2005), Ontologie de domaine pour la modélisation du contexte en recherché d'information. Thèse de doctorat, Université Paul Sabatier, Toulouse, p63.
[15] K. Dubost « Ontologie, thésaurus, taxonomie et Web sémantique » http://www.la-grange.net/2004/03/19.html
[16] O. Corby « Modélisation des Connaissances et Web sémantique : Ontologie », INRIA, Sophia Antipolis, cours essai 2003-2004 http://www-sop.inria.fr/acacia/cours/essi2004/index.html/ontologie2.ppt
[17] M. Angela Biasiotti ,M. Fern_andez-Barrera «Enriching thesauri with ontological information: Eurovoc thesaurus and DALOS domain ontology of consumer law », 3rd Workshop on Legal Ontologies and Artificial Intelligence Techniques joint with, LOAIT '09,Barcelona, Spain, 2009
[18] http://protege.stanford.edu/
[19] http://jena.sourceforge.net