Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 191 (2015) 1266 – 1274

Procedia
Social and Behavioral Sciences

WCES 2014

Implementation and Evaluation of an Indexing Model of Teaching and Learning Resources

Kamal El Guemmat[a]*, El Habib Ben LAHMAR [B], Mohamed Talea[a], El Khadir Lamrani [b]

[a] *Université Hassan II - Mohammedia – Casablanca, Faculté des Sciences Ben M'sik, Laboratoire de Traitement de l'Information, Cdt Driss El Harti, BP 7955 Sidi Othman Casablanca, Maroc*

[b] *Université Hassan II - Mohammedia – Casablanca, Faculté des Sciences Ben M'sik, Laboratoire de Technologie de l'Information et Modélisation, Cdt Driss El Harti, BP 7955 Sidi Othman Casablanca, Maroc*

Abstract

With the advent of teaching and learning resources (TLR), indexing becomes essential to ensure his identification, adaptation, reuse and sharing. Several indexing models of TLR emerged, the problem is to go on a coherent model, in this article we designed our previous contribution of a model entitled MIMTLR (El Guemmat et al, 2013a), which defines a Multi indexing model of Teaching and Learning Resources that aims to enhance the limited indexing LOM standard, with a semantic content indexing is based on ontology's. The purpose of this paper is to evaluate, implemented of MIMTLR and validate it by a programming languages, to ensure that it best meets the constraints provide by a powerful model indexing of TLR. We will present simulation and the advantages of this model that will be useful for the involved in information and communication technology for teaching and learning (ICTTL) especially E-Learning.

Selection and peer-review under responsibility of the Organizing Committee of WCES 2014

Keywords: TLR, Indexing, LOM, Ontology, Validate, E-Learning

1. Introduction

With the arrival of indexing and information retrieval systems, actors of ICTTL invested a lot of effort to end out with indexation models adapted to areas of the ICTTL. This indexing can bring many benefits we cite identification, adaptation, reuse and sharing of TLR.

* Kamal El Guemmat. Tel.: +212-0614-55-71-42.
E-mail address: k.elguemmat@gmail.com

Selection and peer-review under responsibility of the Organizing Committee of WCES 2014
doi:10.1016/j.sbspro.2015.04.654

Several Models indexing of TLR emerged, the problem is to go on a coherent model, implement it, test it in a semantic search engine adapts to this fact. TLR can be represented at two levels, economic, technical, educational and at a level that describes the semantics of the content level. The representation of the first level is in most cases made by description schemes of TLR while the representation of the second level can be done either with external resources, digital dictionaries, ontologies (El Guemmat et al., 2012a), thesauri ... or with internal resources including body, but most compelling is that of ontology which ensures best formalization of domain knowledge in the form of concepts and relations between them may represent the semantic content of a resource. On the other hand it offers a common and shared understanding of a domain, as well as human users and at the level of software applications (Amardeilh, 2007). The purpose of this paper is to implement and evaluate the Model MIMTLR (El Guemmat et al, 2013a) that brings these different levels (economic, technical, educational, semantic content) in a more powerful ontological model. For the implementation we proposed an application for validation we tried to ensure that the model best meets the constraints ensure by a model indexing of TLR. In this article, we define the issues of indexing of TLR in Section 2, we identify The indexing by description schemes of TLR in Section 2.1, The classic indexing in Section 2.2 the more that can provide semantic indexing content in the section 2.3, section 2.4 presents a combination of indexing by description schemes and semantic indexing of content, our contribution in the field will be presented in Section 3 to be divided into two sections, the first 3.1 which shows the implementation of MIMTLR the second section 3.2 that validate the MIMTLR, Section 4 presents a discussion, Section 5 is the last for conclude, present the contributions that will be presented by our model in ICTTL at the end define our perspectives.

2. Model of TLR Indexing

2.1. The indexing by description schemes of TLR

The description schemes consider a description language for digital education resources; the idea is to index very quickly to accelerate the task of information retrieval.

We have several description schemes Standard (LOM, SCORM, IMS-LD ...) but we opted to LOM.

- LOM (Learning Object Model)

Metadata model developed in 2002 by the IEEE consortium for describing resources for educational purposes.

It ensures the treatment, evaluation, sharing exchanging and reuse of digital learning resources.

It is organized into nine categories:

- General: context-independent (id, title, language .....)
- Lifecycle: History and Evolution (Versions ....)
- Meta-metadata: Description of Metadata Features (languages, contribution .....)
- Technical: Format, size.....
- Educational: Educational and pedagogical Characteristic Right: Copyright…….
- Relation: Relation between LO.
- Annotation: Comment.....
- Classification: Relationship with classification systems.

Whenever we can instantiate the model in a particular context is called application profiles, to reach a new profile must interpret, refine, spread, simplifying the syntax and semantics of LOM.

2.2. Classic indexing

Indexing can be done by different means:

- Manual—The document is analyzed by a human expert in the field;
- Automatic—A fully automated process;
- Semi-automatic—It is based primarily on automatic mode. However the final choice remains to the expert in the field to select the significant terms.

The automatic indexing mode provides benefits, given its ability to automate the indexing process.

It includes several treatments on the documents: Automatic extraction of descriptors, the use of an anti-dictionary to remove function words, stemming, the identification of groups of words, the weighting of words before creating the index.

This set of weighted terms will be used to form are presentation of the contents of the document, these terms are organized into a representation It depends on the model of IR (Information Retrieval) that we use (Boolean models, vector models, probabilistic models) (BAZIZ, 2005).

Among the problems that confront the traditional indexing, the ambiguity of words and their disparity.

Textual entities that represent TLR and queries are specified by keywords from the content (Fatiha, 2008).

The ambiguity of words, called lexical ambiguity, refers to words and lexically identical with different grammatical functions, it is generally divided into two types, the syntactic ambiguity and semantic ambiguity.

The disparity of words (word mismatch) refers to different lexical words have the same grammatical function.

Various solutions are proposed to overcome the limitations of traditional indexing.

As a solution to the problem of the ambiguity of words, is to use compound expressions (Fagan, 1987) (Salton, 1988), to reduce ambiguity.

Yet it is not always possible to provide a compound expression in the query that meets the desired direction, and the formulation of expressions requires a great effort from the user.

Solution to the problem of disparity of words is to expand the query using a thesaurus of synonyms (Salton, 1983).

To add a word in the query by its synonyms, we must not only know the word in the query, but also the word that is used to extend (Krovetz, 1992).

As part of TLR Indexing, A new type of indexing appeared to overcome the limitations of traditional indexing and limitations of indexing standards, called semantic indexing, which will be explained in the next section.

2.3. Semantic indexing models of the content of TLR

The semantic indexing provides outcomes at the representation of documents and queries. This is a specialization of traditional indexing, according to (Amardeilh, 2007), the goal is to index by the meaning of words rather than by words. In a context where the ambiguity is present, the semantic indexing is meant to improve the performance of IR.

However there are several semantic indexing model adopted in the case of TLR, we will quote some one:

- Model of thematic description (Hernandez et al., 2008) that index the TLR by ontology concepts.
  The ontology represents a Training Modules and concepts that will be given to students for each module.
- SOAF (Ana Cernea et al., 2008) is a system that provides a model for the automatic extraction of semantic descriptions of multimedia learning resources based on latent semantic indexing (Pecenovic , 1997) using the representation of the resources of the vector space Rn by their characteristics visual, the system allows a semantic description is based on the collaborative user tags.
- Model of semantic indexing over a new algorithm based on Ontology OntOAlgo (El Guemmat et al., 2013b, El Guemmat et al., 2012c) and the Model of semantic indexing of computer networks documents based on a new OntONet Ontology (El Guemmat et al, 2012b) which are organized around the following concepts :
  - Notions of application (NoA): are the key concepts that models the TLR.
  - Notions of use (NoU): Turns are markers that model a set of elements of knowledge to describe the use case of NoA;

At the end of this section we note that he is tedious to represent an object by the Metadata, representation by Ontology become ideal for the task of searching for TLR (Wang, 2008).

2.4. Combination of the models of TLR description.

A combination of the models becomes a necessity to reap its benefits and to include them in a powerful model that can meet the requirements of TLR indexing.

We found several combinations on a few quotes:

- Model of Multi Faceted representation enriched by semantic content (Hernandez et al., 2008) :

- Model of Multi Faceted representation enriched by semantic content Model of semantic indexing which is based on the collaborative user tags (Ana Cernea et al., 2008).
- The 'MIMTLR'. This model, see Fig. 1, must meet the requirements and needs of e-Learning actors and should therefore take into account two layers (Layer 1: Structure of the object, Layer 2: Semantic content of the object) to generate an index that represents the characteristics Enriched by the TLR semantic content.

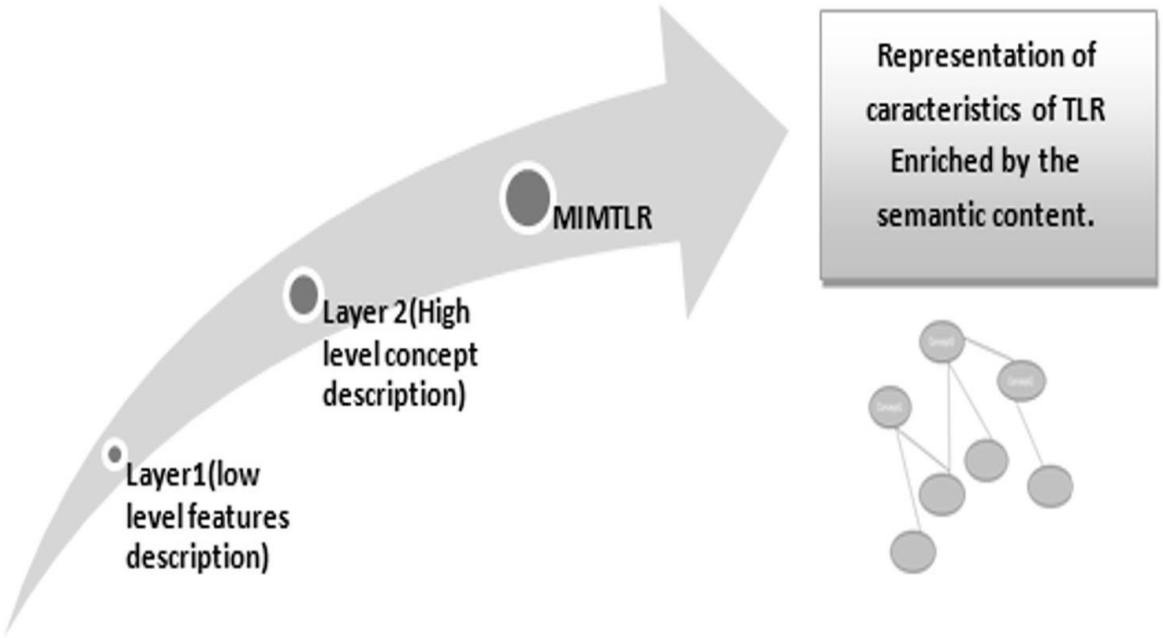

Fig 1. MIMTLR

3. Results

The MIMTLR is a more powerful ontological model ensures these constraints with two layers; it must be implemented and evaluated to ensure its perfect functionality.

3.1. Implementation

In the figure below we provide an overview on the implementation of MIMTLR which is interpreted by the J2EE technology in the Eclipse editor.

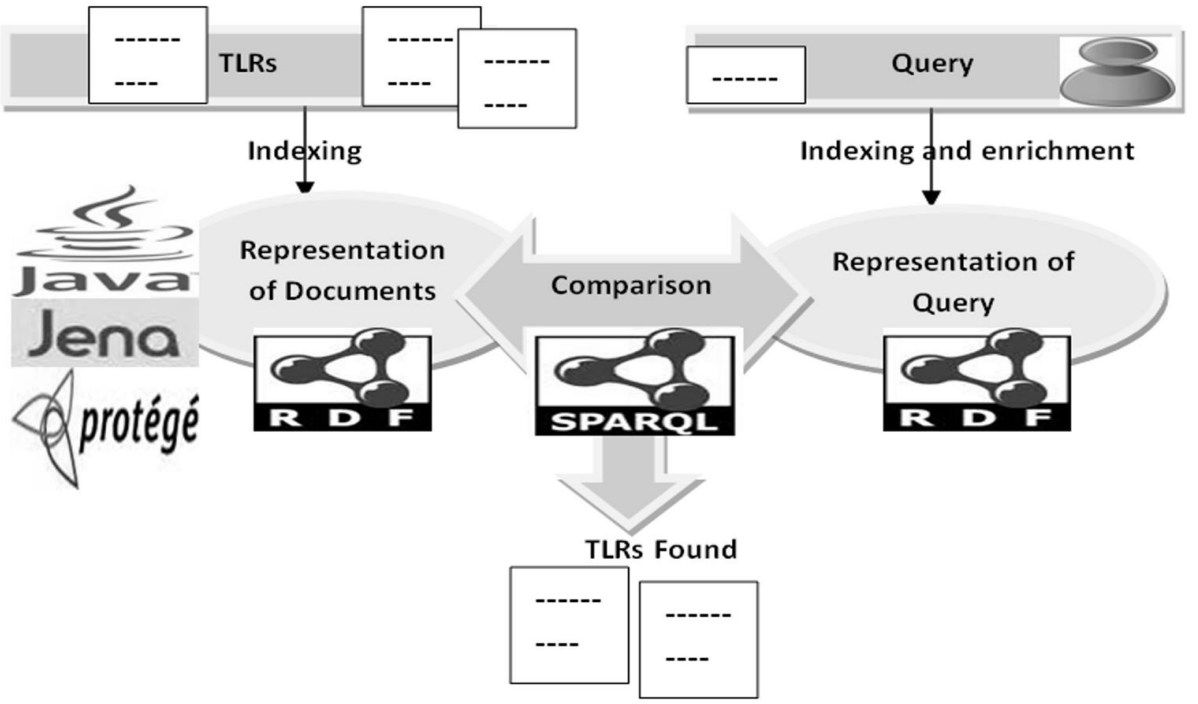

Fig 2.Implementation of MIMTLR

We draw on the general architecture of a system of information retrieval:

- **Indexing Process:**
  - Query: represented by a semantic network in the language of the W3C (please visit : http://www.w3.org/), is RDF (Resource Description Framework).
  - TLR: The recording of information about TLR index on two steps:

Step 1: Indexer will manually index the technical part of TLR.

Step 2: The teacher indexes TLR semi-automatically according to its content, it is based on a predefined ontology and extraction patterns interpreted in a well known java api is Jena, these patterns will extract the NoA and NoU of TLR in this form *NoU*NoA* or *NoA*NoU*, the NoU will be limited between two points (phrase), for more information please refer to (EL GUEMMAT et al., 2013b) for example, the algotithmics domain was seven NoA (variable, test, boucle, tableau, fonction, procédure, tri) and four NoU (définition, syntaxe, types, exemple), then we'll put these concepts and more fields of the first layer in the ontological representation of our model in protégé tool and the TLR will be indexed by a semantic network in the language of the W3C, the RDF. In the end it validates two stages.

- **Research process:**

This is a comparison according to the SPARQL (please visit : http://www.w3.org/TR/rdf-sparql-query/) language to find relevant TLR according queries free language on the structure or content of TLR.

The application we will return the URL of the TLR defined in layer 1, with other information from the layer 1 as additional information on the TLR.

For more technical details, Our application have two servlets (traitement) and (traitementsparql), the first complete the first stage of the indexing process in the homepage indexer filled fields of LOM and indicates the location of the TLR, valid the page, this servlet contains patterns of step 2 of the indexing process.

Then it inserts the fields filled in the ontology and NoUs, NoAs detected in the ontological model MIMTLR, see Fig. 3, implemented in Protégé (please visit : http://protege.stanford.edu/).

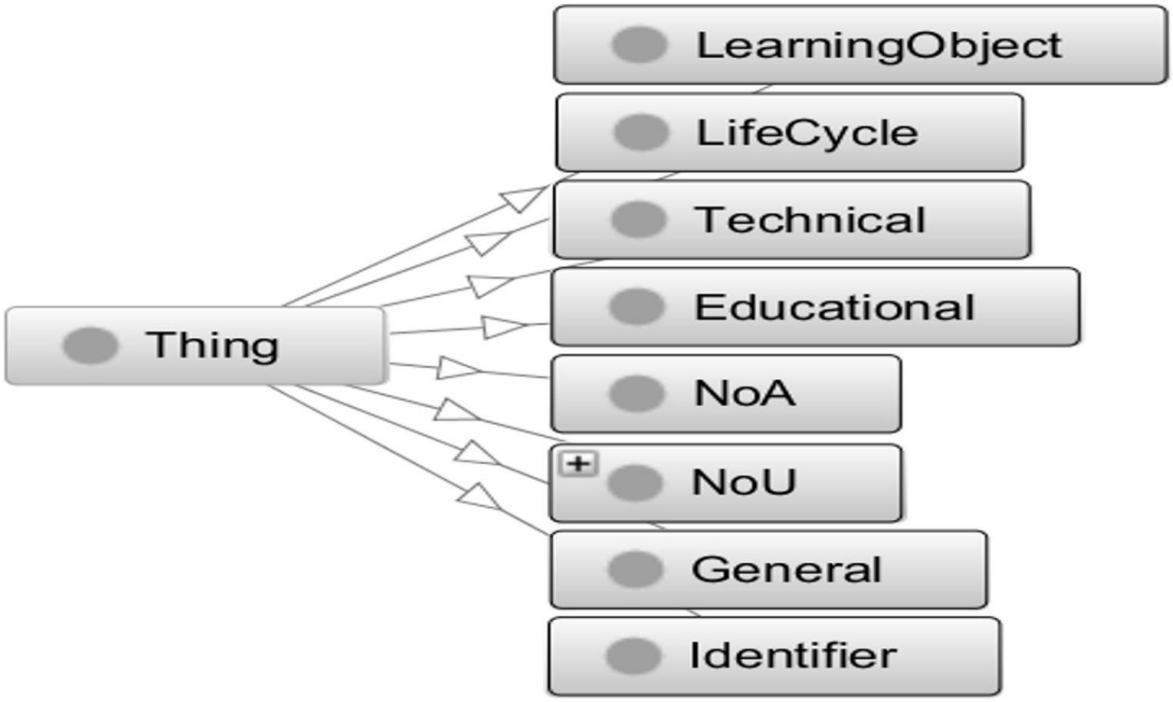

Fig 3 .MIMTLR implemented in Protégé

The second servlet concern step of SPARQL comparison of the research process once the student written a request in free language in sparql.html, see Fig. 4, the servlet returns the best TLR with additional information.

For example we typed query (algorithmique):

Fig 4.Query page

We worked about algorithmics courses, details will be in the next section.

3.2. Validation

We validated our model according two criteria, which are used to evaluate a system for information retrieval:

- Accuracy: Give the proportion of relevant documents to noise ratio (documents not apply)
- Reminder: Give the proportion of documents relevant to silence (documents not found)
- F-Mesure: Give the proportion that combines the accuracy and reminder

A comparison with another model of TLR indexing is necessary to ensure the usefulness of our model.

We compare it with the model indexing implement in the most common Platform for learning MOODLE.

According to our evaluation of the platform, the indexing is based on a classical indexing by Lucenesystem (please visit : http://lucene.apache.org//), with a basic description by schemas description.

It has been 10 requests for the two systems over 13 algorithmics courses (ALGORITHMIQUE ET PROGRAMMATION NON-MATHEUX COURS COMPLET avec exercices, corrigés et citations philosophiques) of Mr. Christophe Darmangeat (please visit : http://www.pise.info/algo/index.htm/), the attached tables interpret the statistics for 10 queries.

TABLE 1. STATISTICS OF MOODLE

Request	MOODLE					
	Exact by expert	Noise	Silence	Accuracy MOODLE	Reminder MOODLE	F Mesure MOODLE
Définition de Variable	1	12	0	0,076923077	1	0,142857143
Variable	2	10	0	0,166666667	1	0,285714286
Boucle	1	6	0	0,142857143	1	0,25
Algorithmique	1	9	0	0,1	1	0,181818182
Fonction	2	4	0	0,333333333	1	0,5
Exemple de fonction	2	11	0	0,153846154	1	0,266666667
Structure d'une boucle	1	12	0	0,076923077	1	0,142857143
Cours édités en France	13	0	3	0	0,8125	0
Leçons destinées aux étudiants	13	0	13	0	0,5	0
Cours pour la range âge entre 18 et plus	13	0	3	0	0,8125	0

TABLE 2. STATISTICS OF MIMTLR

Request	MIMTLR					
	Exact by expert	Noise	Silence	Accuracy MIMTLR	Reminder MIMTLR	F Mesure MIMTLR
Définition de Variable	1	7	0	0,125	1	0,222222222
Variable	2	7	0	0,222222222	1	0,363636364
Boucle	1	2	0	0,333333333	1	0,5
Algorithmique	1	1	0	0,5	1	0,666666667
Fonction	2	2	0	0,5	1	0,666666667
Exemple de fonction	2	1	0	0,666666667	1	0,8
Structure d'une boucle	1	1	0	0,5	1	0,666666667
Cours édités en France	13	0	0	1	1	1
Leçons destinées aux étudiants	13	0	0	1	1	1
Cours pour la range âge entre 18 et plus	13	0	0	1	1	1

4. DISCUSSION

Based on this comparison we concluded that Moodle is not interested in the semantic part especially with regard to the request that hide more semantics for example:

- Cours édités en France: MIMTLR has a field in layer 1 is Coverage which contains information about the place where the course is published.
- Leçons destinées aux professeurs : cache synonymous of 'leçon' is 'cours', which is filled in layer1 in Learning Resource Type knowing that our model once it inserts a field in protégé it inserts its synonyms, it is a kind of enrichment query(Rachdi et al., 2012) .
- Définition de variable: MOODLE gives us any course talking definition or variable against MIMTLR provides courses that address the variable definition.
- Cours pour la range d'âge entre 18 et plus: hides a high level of semantics that is described in our model in layer 1 in the Typical Age Range field.

Even if that differs from our model compared to other model is that NoA concepts and NoU of second layers add the semantic aspect for LOM Category in terms of use of NoA which is not present in other models that are interested only in concepts interpreted by the course forgetting use case of NoA more than the other models do not represent as an ontological view. We introduced the concept NoU to return to the students the courses dealing a NOA in detail, in order to avoid the courses that cover a brief way or namesake of NoA, and enriching the request by Synonyms of NoA and NoU. That way we will have a full technical, educational and semantic indexing of the TLR.

5. Conclusion et Perspectives

According to the implementation of our model we concluded that it is completely valid offers a semantic representation of TLR missing for online learning platforms and add a level of use for notion of courses. Our model will be useful for players of information and communication technology for teaching and learning to index the TLR and to find them and ensure interoperability between them. Our result in multiple perspectives, it is first necessary to find a method of fluid automatic instantiation of the appropriate model and to consider other practical effects of this model on the field of digital learning, we evaluated the identification of TLR but we can evaluate adaptation, reuse and sharing of TLR according MIMTLR model.

References

Amardeilh, F. (2007). Web Sémantique Et Informatique Linguistique : Propositions Méthodologiques Et Réalisation D'une Plateforme Logicielle, Doctorat De Universite Paris X – Nanterre.

Ana Cernea,D.,& Del Moral,E.,& E.Labra Gayo,J .(2008). "SOAF: Semantic Indexing System Based On Collaborative Tagging", Interdisciplinary *Journal Of E-Learning And Learning Objects*, *Volume 4*.

BAZIZ,M. (2005). Indexation Conceptuelle Guidee Par Ontologie Pour La Recherche D'information, Doctorat De Institut De Recherche En Informatique De Toulouse, 2005.

El Guemmat, K.,& Ben Lahmar,E.H.,& Talea,M.,& Chara,A.,& Rachdi,M. (2012b). Indexation Semantique Des Documents De Reseaux Informatiques , 7ème Rencontre Nationale Des Jeunes Chercheurs En Physique – 20 & 21 Décembre 2012 Faculté Des Sciences Ben M'sik, Département De Physique, Casablanca.

El Guemmat, K.,& Ben Lahmar,E.H.,& Talea,M.,& Chara,A.,& Rachdi,M. (2013a). Multi Layers Indexing Model Of Teaching And Learning Resources, International Conference On Engineering Education And Research 2013

El Guemmat, K.,& Ben Lahmar,E.H.,& Talea,M.,& Chara,A.,& Rachdi,M. (2013b). Semantic Indexing Of Algorithms Courses Based On A New Ontology, *Computer Technology And Application, Vol. 4*, No. 3, 2013

El Guemmat, K.,& Ben Lahmar,E.H.,& Talea,M.,& Chara,A.,& Rachdi,M. (2012c). "Construction Of A New Ontology For The Representation Of Algorithmics Knowledges", International *Journal Of Science And Advanced Technology, Volume 2* No 9 September 2012.

El Guemmat,K.,& Ben Lahmar,E.H.,& Talea,M.,& Chara,A.,& Rachdi,M.(2012a)."Etat De L'art : Construction D'ontologie", 7ème Rencontre Nationale Des Jeunes Chercheurs En Physique – 20 & 21 Décembre 2012 Faculté Des Sciences Ben M'sik, Département De Physique, Casablanca.

Fagan, J.L. (1987). Experiments In Automatic Phrase Indexing For Document Retrieval: A Comparison Of Syntactic And Non Syntactic Methods, Ph.D. Thesis, Dept. Of Computer Science, Cornell University, Sept. 1987.

Fatiha, B.A.(2008). Contribution À La Définition De Modèles De Recherche D'information Flexibles Basés Sur Les CP-Nets, Doctorat De L'Université De Toulouse.

Hernandez, N.,& Mothe,J.,& Ralalason,B.,& Ramamonjisoa,B.,& Stolf,P., (2008). "A Model To Represent The Facets Of Learning Objects", Interdisciplinary *Journal Of E-Learning And Learning Objects, Volume 4,* 2008.

Krovetz,R.,& Croft,W.B. (1992). Lexical Ambiguity And Information Retrieval, *ACM Transactions On Information Systems 10* (2) (1992) 115-141.

Linguistics (ACL), Department Of Computer Science, Cornell University, Ithaca, New York, Pp. 204-210.

Ouafia, G.,& Abel,M.H.,& Moulin,C. (2009). "Une Ontologie Pour L'indexation D'objets Pédagogiques", AFIA Platform Workshop: Constructions D'ontologies : Vers Un Guide De Bonnes Pratiques, Hammamet, Tunisie, May, 25-29.

Pecenovic, Z. (1997). "Image Retrieval Using Latent Semantic Indexing". Final Year Graduate Thesis, Audio-Visual Communications Lab, Ecolepolytechnique Federale De Lausanne, Switzerland, June 1997.

Rachdi, M.,& Benlahmar,E.H.,& Elguemmat,K.,& Chara,A.,& Labriji,E.L. (2012). "Enrichment Of Queries For Education Domain", 13th International Arab Conference On Information Technology – ACIT'2012.

Salton,G. (1988). Syntactic Approaches To Automatic Book Indexing, In: Proceedings Of The Annual Meeting On Association For Computational

Salton,G.,& Fox,E.,& Wu,H.(1983). Extended Boolean Information Retrieval, *Communications Of The ACM 26* (11) 10224 1036.

Wang, S. (2008). "Ontology Of Learning Objects Repository For Pedagogical Knowledge Sharing", Interdisciplinary *Journal Of E Learning And Learning Objects, Volume 4,* 2008.