Intelligent Agents for the Semantic Annotation of Educational Resources

Aziz ORICHE, Abderrahman CHEKRY, Mohamed KHALDI

Abstract- Our objective is to describe the content of educational resources semantically annotating with unambiguous information to facilitate the exploitation of these resources by software agents, these resources are delimited by tags (XHTML, XML), and are well structured. We propose a semantic annotation system based on three intelligent agents to manage semantic annotations educational resources, these annotations are guided by domain ontology. The domain ontology contains a set of concepts, relationships between these concepts and their properties. Each concept of ontology some have one or more (synonyms) to represent words. All these terms are used to describe instances of a domain concept; these terms are more focused in the description of concepts. All these concepts are validated by the domain expert. Taking into account that the teaching materials are in HTML or XML whose structure is a DOM (Document Object Model) we seek to identify the terms of educational concepts in the nodes of the tree, extract all these concepts while respect of the domain ontology. The terms or educational concepts candidates are associated with terms and concepts of the ontology to determine appropriate concepts to annotate nodes in which they are located.

Index Terms: Semantic Annotation, Metadata, Multi-agent systems, Ontology

I. INTRODUCTION

Generally many documents and resources can be used as part of an e-learning training (or e-Learning). Some of these resources are produced "in house" by the various factors involved in the training; others are available on the web: online courses, course materials, supports oral presentations (slides), bibliographies, frequently asked questions, lecture notes, etc. [1]

To meet the ever-increasing number of resources, the search engines should be able to provide more detailed answers and handle more complex queries integrating knowledge of the user. Given the need for a formal description of the content of educational materials online, we see that it is interesting to integrate software agents able to label semantically textual and multimedia resources, metadata, on the one hand, and to exploit and improve information retrieval, on the other hand. The formalization of existing web pages is defined as the task of semantic annotation and automatic.

The automatic existing annotation systems are essentially syntactic. Automatic annotation of semantic aspects of web pages requires solving the problem of identifying the appropriate concepts to the content of each web resource and the profile of each user type, vis-à-vis ontologies.

Manuscript Received November, 2013.

Aziz Oriche, Université Abdelmalek Essaâdi, Laboratoire, LIROSA & LNTE, Tétouan, Maroc.

Abderrahman Chekry, Université Abdelmalek Essaâdi, Laboratoire, LIROSA & LNTE, Tétouan, Maroc.

MOHAMED KHALDI, Université Abdelmalek Essaâdi, Laboratoire, LIROSA & LNTE, Tétouan, Maroc.

Current data from the web are usually written in natural language, as for humans. Natural language is inherently ambiguous too; formal and explicit semantic alternatives should be implemented to eliminate the ambiguities of natural language both in the content of the resources and in their annotations. The task of annotation for the semantic web is therefore to take as input a resource document and outputting the same document enriched with semantic annotations based on representations of knowledge more or less formal content.

In this article we will solve problems related to: (i) the extraction of terms and concepts teaching candidates, (ii) combine the concepts extracted ontology, (iii) the annotation that is tailored to the content of node.

We present in this article the role of ontologies for structuring educational resources. We specify our approach to semantic annotations of educational concepts by intelligent agents finally we conclude.

II. THE SEMANTIC WEB

The Semantic Web is a powerful environment for implementing e-learning applications; it is a technology in which machines will think clearly on a semantic basis, it is characterized by the sharing of knowledge based on ontology structure.

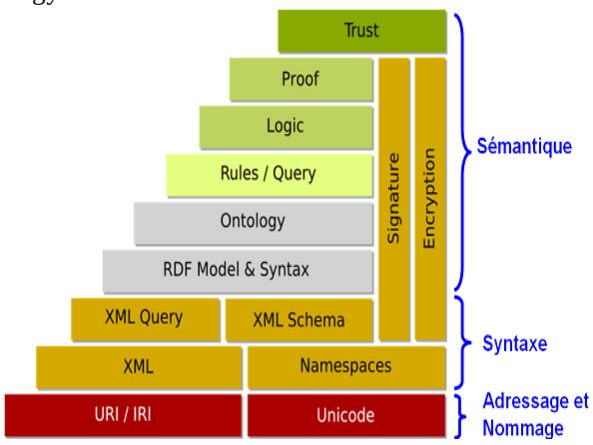

Figure 1: The layers of the Semantic Web

The idea of te Semantic Web is to model the content of Web resources by adding semantics as metadata to make resources understandable by machines, give a formal representation to semantic knowledge base so that it is understandable by programs and finally Structure the web content to enable machines to automatically process this content and do reasoning about the represented knowledge.

A basic language of the Semantic Web (Fig. 1) is the XML markup data and RDF (Fig. 2) to represent the data that machines can do useful reasoning tasks on documents. OWL is an

ontology language for the Web. This is part of a scalable stack relative to the W3C Semantic Web recommendations.

The ontology and semantic annotation play an important role in the Semantic Web. Ontologies are necessary for semantic representations requiring a community consensus. Semantic annotations are to add structured metadata to information resources on the Web.

A. E-learning

In June 2000 the European Commission defines e-learning as "the use of new multimedia technologies and the Internet to improve the quality of learning by facilitating access to resources and services, as well as trade and remote collaboration. »

In many definitions of e-learning that can be encountered, the focus is usually on the use of the network and the provision of educational content (as opposed to a linear course). The trainer's role is to assist the learner instead of transmitting his knowledge. A logical transmission passes to a logical accompaniment; the pedagogy of Trainer is based on the personal work of the student. Therefore it has an active role in his formation and content made available cannot be a course outright online. Educational resources should be described in greater detail; it is for the learner to access these resources and trainer to guide him to access.

The content of an e-Learning training is usually divided into granules or grains of knowledge. The determination of these granules is a complex task that cannot be achieved only by instructors or tutors, but by a set of experts (editorial board), including in particular the head of training considered.

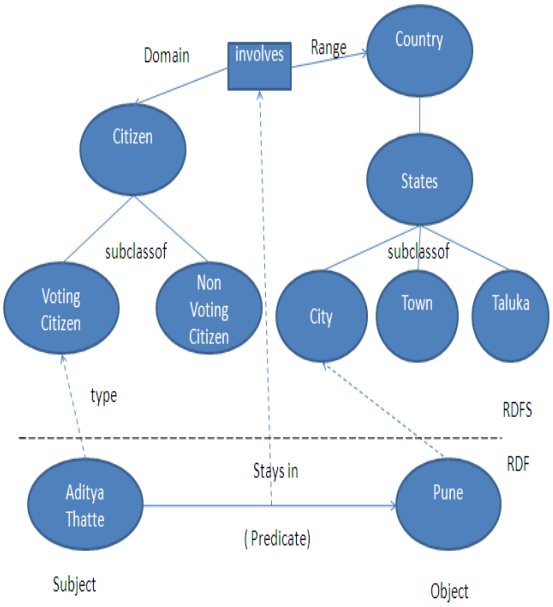

Figure 2: Example RDF schema

Among the many documents found on the web, we are interested only documents known teaching, which are created or that can be used for learning. It is therefore essential to organize and index.

Educational materials become learning objects: digital entity used in the learning environment, with the following properties:

- Autonomy: Each learning object can be used independently of the other;
- Reusability: An elementary learning object can be used in different contexts and for multiple purposes;
- Aggregation: Learning objects can be combined to form other learning objects;
- Indexing: Each learning object has a description for easy retrieval. In addition, a learning object should lead to learning:
- Without superfluous: If only a fraction of the course is required, it must be limited to relevant learning objects;
- In time: With indexing learning objects we can find them instantly;
- Custom: Learning objects allow customization of courses across an organization or individual.

B. Semantic Web and e-learning

The key to the architecture of the Semantic Web property is offered by an appropriate set of agents appears powerful enough to meet the requirements of e-Learning system: fast, just-in-time and relevant learning. The e-Learning material is semantically annotated for a new learning demands; it may be easily combined in a new learning course [2].

According to their preferences, a user can easily find useful learning content. The process is based on applications of the Semantic Web and navigation through the learning material activated by an ontological skeleton.

Thus, the Semantic Web can be treated as a suitable platform for implementing an e-learning system, as long as it provides all the means for ontology development (learning) [3]. However, the content is useless unless it can be easily searched and indexed. This is true when the volume and type of content increases. The e-Learning community has agreed to use metadata. Metadata provides a common set of tags that can be applied to any resource, regardless of who created it and what are the tools that will be used and even where they are stored. Tags are data that describes other data. The metadata tagging enables organizations to describe, index and search their resources [4].

Thus, we believe the approach of a semantic web appears as the most relevant for our project. Indeed, this approach emphasizes the problems of indexing and semantic inference to those annotations. In a learning environment, the user will gain navigate through representations offered (ontologies, semantic annotation). These representations obtained from a formal language understandable by any type of user, they serve as a kind of metadata. In this way, a user can access the same document in multiple points of view. Users then develop their own opinion.

III. ONTOLOGIES

A. Definitions

Ontology includes or implies a certain view of the world in relation to a given area. This view is often conceived as a set of concepts - entities, attributes, processes - their definitions and their interrelations. This is called a conceptualization.

The ontology enables the organization of learning material to turn small pieces of semantically annotated learning objects. Items can be easily organized into learning course (fast and just in time) and delivered on demand to the user according to his profile and business needs (relevance) [5].

In the Semantic Web, the ontology allows the user during a Web search to access not only to documents related to keywords in the query, but also those that are related ontologically (semantically) to them, this makes the search more relevant. It aims to describe concepts and relationships that bind them, and with deduction rules to make them more understandable and usable by the different agents (human or software) [6].

Ontologies are so central to the Semantic Web, which, on the one hand, seeks to rely on models of the Web resources from conceptual representations of the areas concerned and, on the other hand, aims to allow programs make inferences above. Research on them is therefore essential.

Different types of ontologies are distinguished: domain ontologies specific to a particular area and sometimes reusable ontologies of the application, which contain the elements necessary for a given application, generic ontologies that cover several areas, and finally meta-ontologies that describe primitives for languages "representation" of knowledge.

B. Ontologies and e-learning

Ontologies are used in e-learning systems:

- Allow each member of the organization to understand the meaning of each term used;
- Foster communication between all members of the organization, with a clear and meaningful definition of all concepts.
- Better indexing information resources described ontological concepts.
- Definition of a common conceptual vocabulary of the domain knowledge
- ontologies act as amplifiers of intelligence
- The ontological relationships to structure all of these concepts by the links they establish and the attached to each semantics;
- The structure thus created fosters retention of information by placing each concept relative to those that are related to it.
- Ontologies lend them to concrete and visual representation that helps the learner to navigate easily within knowledge.

IV. THE SEMANTIC ANNOTATIONS

The semantic annotation is to assign an entity (a string, a sentence, a paragraph, part of a record or document) a metadata whose semantics are often defined in a model. This metadata can be stored in the document itself, or in another document referencing the entity annotated by URI (Universal Resource Identifier). The external annotation is more interesting because it allows you to reference once a metadata with entities in several documents. In this case, an annotation is a set of instances (metadata) related to Web resources. The annotation of documents is a process similar to the process of information retrieval. The difference is that the extraction does not preserve the context of appearance (the document) of the entity.

The Semantic annotation of educational resources with metadata allows a better interpretation of their content. When these annotations are available, it is possible to formulate queries based on ontology vocabulary, reasoning about these annotations, and collect responses combining data from different documents. Search engines are then able to go beyond search keywords and accurately answer complex queries.

One of the main difficulties lies in the fact that educational resources are very heterogeneous both in terms of format (HTML, PDF, GIF, AVI, MPEG…) and their structure from the point of view of the vocabulary used. The approaches proposed in the literature concerning complementary fields such as data mining, artificial intelligence, knowledge engineering and natural language processing.

In order to reduce the problems associated with the annotation, the developed methods often focus on an application domain, a class of web sites or on specific entities.

A. The methods of semantic annotations

An annotation method is characterized by the level of structuring documents it processes. Some approaches are specialized in the annotation of free text and other elements in the annotation of specific structures such as tables and lists. However, existing documents are often structurally heterogeneous and include well-structured parts and textual parts.

Many extraction and annotation approaches of semi-structured documents exist in the literature. Some aim to annotate documents, others to build a knowledge base (a populated ontology instances), or both. These approaches differ not only in their purposes but also by the assumptions they operate, and are often tailored to specific corpus. These assumptions may involve the use of:

- Approaches exploiting the document structure

The structural organization of the elements in the semi-structured documents can be exploited in an information extraction process. Many approaches exploit the document structure elements such as tables, lists, nesting of these elements to automatically or semi-automatically generate models or patterns based on the structure. These patterns are used to identify concepts, instances of concepts and / or relationships between these bodies.

- Approaches exploiting patterns in the text

The use of lexical and syntactic patterns is to segment a text into entities and to define the different entities. The lexicon used in pattern may come from a dictionary.

- approaches exploiting lexical resources of the web

The aim of the approach is to semantically annotate named entities found in a document that already exist in the knowledge base.

In our approach, the candidate instances are extracted using the extraction agent. To associate a candidate for a concept instance, we use the ontology directly. Moreover, concepts extracted correspond to the context and also associated with concepts of ontology. After detection of known concepts, comprehensive user and corrects the detected elements corresponding to instances of primitive concepts

B. Ontology and semantic annotations

Annotation or metadata refers to particular providing information about a resource. In terms of documentation is secondary information affixed to a primary resource is the document semantic annotation provides more basic information such as the title and the authors a more accurate description of the

knowledge contained in the document and their semantics with respect to the field. A semantic annotation must have a predefined structure and clean the area studied for the free generation of an annotation can create problems of ambiguity in the definition of redundancy and reuse. Furthermore, the use of different structures makes use of these annotations with an engine very difficult semantic search.

Being based on a predefined set of concepts, properties and relations (i.e. is based on ontology) make the semantic annotation a more interesting point of view of the structure and content. This guide enables its annotated document does not confront a problem of ambiguities and know how knowledge contained in the document must annotate.

In [7] the authors emphasize the value of using ontology for the creation of semantic annotations and [8] provides a comparison of the search results of two systems one is based on annotations generated freely and on the other annotations based on an ontology.

However, despite its importance for the management of domain knowledge annotation process is very heavy and requires more resources what motivated several works on automating this task.

V. INTELLIGENT AGENTS

A. Definition

The term "intelligent agent" refers to a number of applications running at the same time in an Internet environment and other environments such as operating systems. The definition is very general. It refers to a part of a computer system (program code, crawler, spider…) or complex system itself (Meta engine online price comparison, software). However, the term covers a wide range of applications.

The multi-agent systems (MAS) provide a new method for analyzing, designer and implement sophisticated applications because they are part of the IAD (Distributed Artificial Intelligence) area is also benefiting from other disciplines such as cognitive science, sociology, and social psychology. Today, most applications require distribution of tasks between autonomous "entities" (or semi-autonomous) to achieve their goals in an optimal way. Since traditional approaches are generally monolithic and their concept of intelligence is centralized, current applications are established based on multi-agent system [9].

A multi-agent system allows a decomposition of complex tasks into subtasks, which reduces the complexity of the system volume level of knowledge and facilitates the development, testing and updating. In addition, a distributed architecture ensures distribution control using the mechanisms of cooperation and coordination. In addition, a multi-agent system allows the integration of newly updated for the treatment of certain linguistic phenomena or to use different methods for the same problem strategies.

B. Intelligent agents in e-learning

Intelligent agents - so-called e-assistants or assistance programs - can sit inside a computer and make learning e-learning dynamically based on the need of user. They can help to extract relevant to the user based on the information available knowledge. Intelligent agents can play a major role in the E-learning based on ontologies. Intelligent agents can customize the content of e-learning to provide the knowledge and skills of the user. They can support the choice of the user in various areas, the level of knowledge and learning style and accordingly providing information dynamically.

A dedicated e-learning can be built with the technology of multi-agent systems that would be responsible for finding the best match of the request based on the needs. The e-learning system based agents would be responsible for the information dynamically adapted as and when requested by the user. Agents could act as search engines and retrieve user information beneficial to the Internet on the basis of which the e-learning module of the user could be made more interesting and interactive. Each agent can be programmed according to the profile of the student.

VI. TOOLS FOR SEMANTIC ANNOTATION

A semantic annotation tool is a software tool that allows you to insert and manage semantic annotations associated with at least a given information resource. Under the Semantic Web, the semantic annotation tools use an ontology, or at least a formal model, which formalizes the structure and annotations produced according to the concepts and constraints defined in this ontology.

These semantic annotation tools aim to ease the burden of manual annotation Resources Management documentaries. Most of them have evolved into environments increasingly automated using the methods from the fields of Information Extraction and Learning Systems [10]. In addition to the traditional manual annotation interface, these methods are able to suggest to the user a set of semantic annotations for the documentary resource analyzed.

A semantic annotation tool can also be used to populate an ontology, i.e. to instantiate the knowledge base containing the bodies of the reference ontology. The Semantic Web approach is principally engaged in the automated production of semantically annotated documents.

A. Study examples of annotation tools

The semantic annotation tools are generally in the context of the Semantic Web. They are used to create and manage annotations of document content. They are based on a formal model of knowledge in general ontology, and exploit more of the Semantic Web standards. Most automatic annotation tools studied: MagPie [11] [12] (Fig. 3), semTag [13] [14] (Fig. 4) and Armadillo [15] [16] [17] (Fig. 5) generate annotations using RDF, XML or DAML + OIL,and use predefined ontologies.

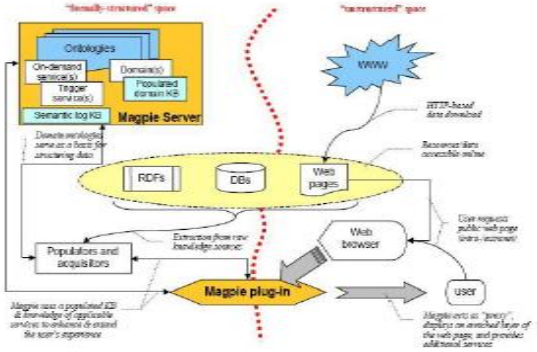

Figure 3: Magpie Architecture Overview

We note that the tools may produce incorrect or empty references. We can therefore

conclude that the semantic annotation tools have *several drawbacks in the separate components of resource materials processing and in the use of information extraction tools. Compared to our goal, these tools are not interested in the description of the whole of an information resource, and more specifically to the development of semantic links between these resources. This description is described by RDF metadata annotation. From the semantic point of view, the annotation process often results in an ontology population with the detection of new instances of concepts or relationships between instances that come to enrich the ontology [18][19].*

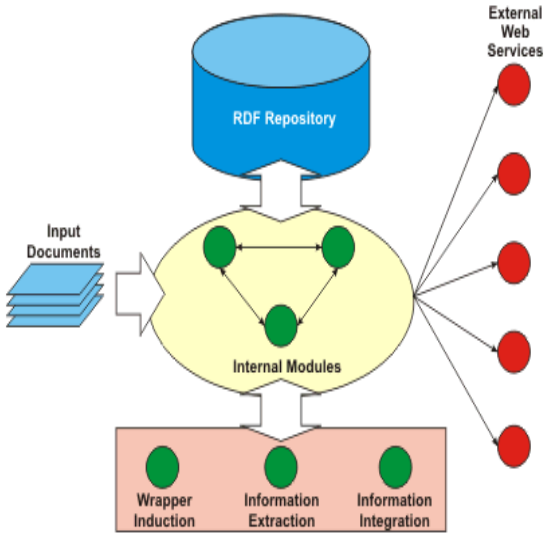

Figure 5: The Armadillo architecture

B. *Synthesis of existing tools*

Under the Semantic Web, it is intrinsically linked to an ontology modeling. Indeed, this ontology will represent the concepts, attributes and relationships of a domain using a knowledge representation language like OWL Web oriented. The semantic annotations are structured using this ontology and values point to instances of the reference ontology, or in some cases directly to the concepts themselves.

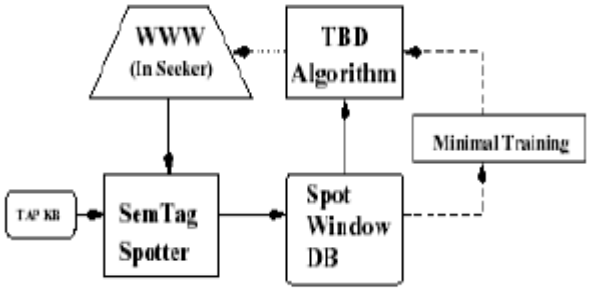

Figure 4: The SemTag architecture

There is a fairly wide range of tools for semantic annotation each having its own characteristics but aimed more oriented towards assisting human annotators to create annotations. Depends on the level of automation of the motor built in the tool information extraction. Annotations can be stored in an annotation server, embarked or disembarked with respect to the original document, or in a knowledge base mode. They can, among other things, be used to improve information retrieval systems, to populate an existing ontology or to assist in the construction of ontologies. We learned from our study that the tools presented are all more or less based on research.

Despite the rapid evolution of languages and standards, thanks to the development of the Semantic Web, these tools are still not suitable for practical applications in the real world. And that e-learning applications benefit from the new model specified by the founder of the Semantic Web framework, some still existing limits need to be pushed.

- Extraction engines are mostly based on supervised learning process.

These systems may be effective for structured or semi-structured content. But for e-learning applications, the content is educational, potentially containing new strategic information, including semantic relations between concepts, and whose semantics is more difficult to extract [20]. I think in this case it is necessary to prioritize based on various fine linguistic analysis to identify the semantic content of the documents processed extraction systems.

- The semantic annotation must continue to focus on approaches based on domain ontologies, not generic. This is especially the case in the applications for the world of Learning.
- It must be able to provide maximum assistance to the user. Processes and interfaces must not only submit suggestions, proposals, but they must also be able to guide in particular taking into account the constraints and modeled in the reference ontology restrictions. I am convinced that the question of the integrity of annotations, especially in the knowledge base that welcomes new instances from these annotations is crucial for the exploitation of the results provided by such application.

C. *Work on semantic annotation*

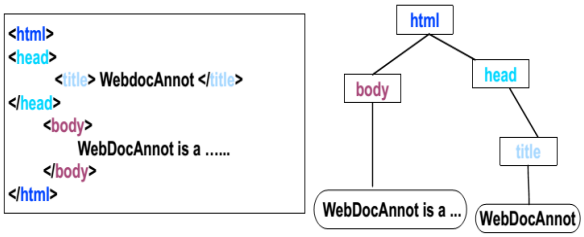

Figure 6: Example of DOM document

We studied the recent approaches to semantic annotation and systems for automatic generation of semantic descriptions such as Framework WebCat [21] and ALLRIGHT system [22]. These works use the standard XML-RDF annotation. Compared to our goal, they are limited to a portion of educational resources (texts, images, etc...) for the extraction of RDF metadata. In addition, they do not address the semantic links between these resources during this extraction. Compared to our goal, they ignore the user profile when learning. Based on these studies, we see that the results of these studies are insufficient with respect to our semantic annotation of learning resources aim after interrogation by the user.

The objective of this work is then to generate metadata for information resources. These metadata are semantic annotation of these resources. Most of these

works are based on an ontology.

We will follow the same approach as the work: annotate resources based on an ontology. But they have several limitations compared to our goal. Indeed, we want to get to provide semantic annotation of learning resources taking into account the links between these resources. This annotation will be more relevant in the case of use of an ontology.

We also suggest that the question of the annotated resources is done by expanding a query language supporting the semantic dimensions of annotation metadata. To achieve our goal, we propose the creation of a new approach to semantic annotation of learning resources. We present in the next section the approach of our proposal.

VII. OUR APPROACH TO ANNOTATION

Directed to a system for automatic generation of semantic annotation for items domain e-learning two points are considered:

What ontology can be constructed based on annotations and what tools will handle text and extract relevant information.

To the extent our first choice was focused on the design of a domain ontology that covers the concepts of our e-learning application.

For annotation, our approach is automatic; we seek to define instances of concepts. These are the nodes of documents, where terms and educational concepts associated with ontology concepts are located. This annotation takes into account the semantics of terms or pedagogical concepts and structures containing nodes.

Indeed, we exploit the structure not only to define the instances but also to detect any relationship between these bodies in the field.

The steps of our method and the general architecture of our system are presented below.

A. The ontology used

In our project, we consider domain ontology is a specification of all relevant concepts in our application. Each concept of this application is the concept of the ontology which are grafted on to other relationships and concepts. This ontology includes training in general and not training for a particular area.

In this context, our ontology contains a set of concepts, relationships between these concepts and their properties. Each concept of ontology some have one or more (synonyms) to represent words. All these terms are used to describe instances of a domain concept; these terms are more focused in the description of concepts. All these concepts are validated by the domain expert.

The teaching materials are in HTML or XML, the structure is a tree DOM (Document Object Model) (Fig. 6) as we seek to identify the terms or concepts in teaching the nodes of the tree, extract all these concepts in the area of respect ontology, these terms or educational concepts candidates are associated with terms and concepts of the ontology to determine appropriate concepts to annotate nodes in which they are located.

B. Semantic annotation system

In our work, we have defined an approach to automate the semantic annotation of educational documents, unsupervised and guided by ontology (Fig. 7).

Our approach identifies terms and concepts in educational nodes in the DOM (Document Object Model) representing the structure of a document (HTML or XML) and these are the nodes that are annotated as instances of concepts. We use the annotated nodes to infer the possible existence of a semantic relationship domain annotate.

We have defined an architecture using the W3C languages RDF (S), OWL and SPARQL composed mainly three intelligent agents each has its own business module: agent extractor educational concepts (AGEC), semantic annotation agent (AGAS) and agent treatment (AGET).

- Agent extractor pedagogical concepts: This agent extracts terms and pedagogical concepts (PC) in the documents under an extraction that we have defined and associated with the concepts of the ontology module. This agent is based on patterns of self-extractions in the field to retrieve educational concepts and terms related to teaching. We will use an algorithm to extract each time a term or concept teaching candidate has been associated with a concept of the ontology.

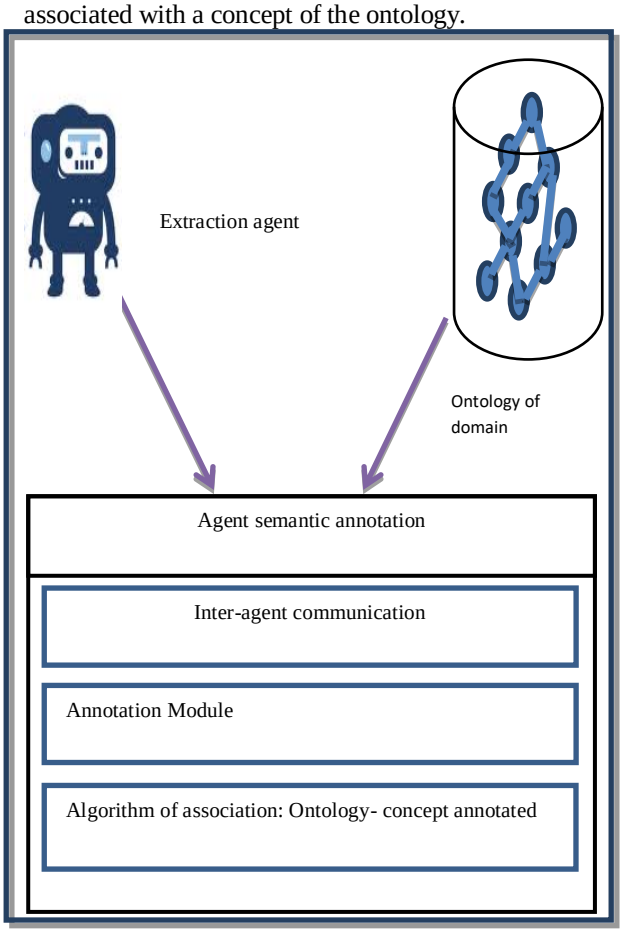

Figure 9: Architecture of annotation agent

Agent semantic annotation: This agent annotates the document nodes in accordance with the annotation module that we have defined. A party of this module is generated automatically from the ontology to represent the extraction results and annotation.

- Treatment Agent: This agent can process queries on the warehouse annotations; the answers to these queries are consistent with a learning module that defines the needs of each user.

VIII. ARCHITECTURE OF OUR SYSTEM

A. *Extraction Process*

The extracting is applied to a set of educational documents after pretreatment. It is based on an extraction module with patterns of extractions distinguishing parsing to extract candidate concepts and relationships used in detection techniques to extract relevant fragments labeled relations between concepts and an algorithm to retrieve each time a term or concept teaching candidate has been associated with a concept of the ontology. The output of this agent (Fig. 8) is a set of RDF metadata, the metadata annotation of pre-specified terms or concepts which PC were identified and in which nodes of the documents

B. *Process annotation*

The semantic annotation agent (Fig. 9) aims to annotate the learning materials based on the result of the extraction agent.

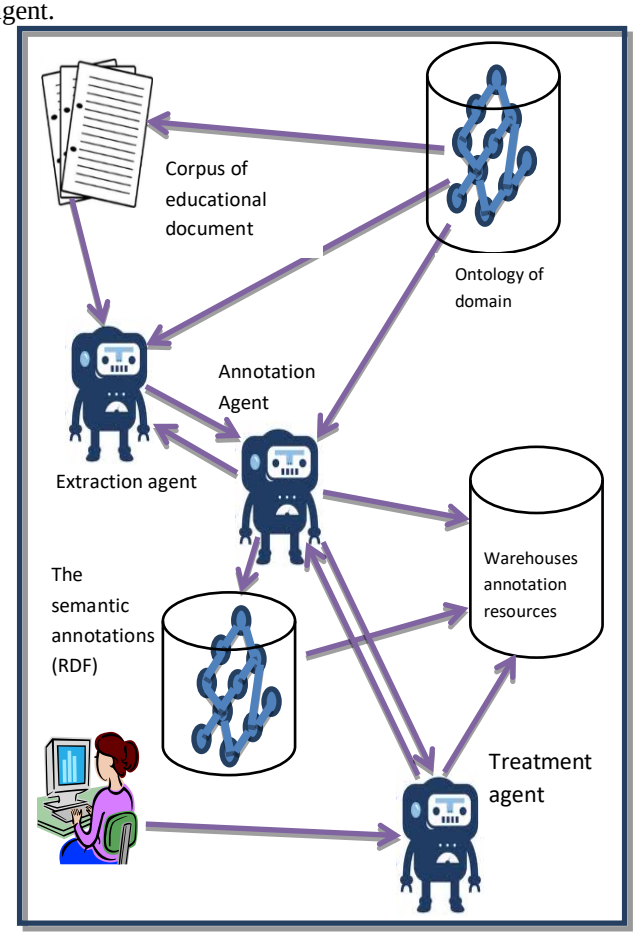

Figure 7: Architecture of our system

To do this, we define an algorithm to generate semantic annotations of module annotation as follows:

- Annotate nodes taking into account both concepts associated with these nodes in the extracting and their structural aggregation in the documents.
- Infer relationships between nodes, taking into account both the types of terms and PC located in these nodes and their neighbors.

C. *Handling Process*

The treatment agent (Fig. 10) is the agent that handles the user request expressed with the concepts and relations of the ontology. The aim is to automatically reformulate the initial query by exploiting the warehouse annotations to achieve the sort of answers and responses based on their semantics.

In our work, we have defined an algorithm to construct reformulated queries

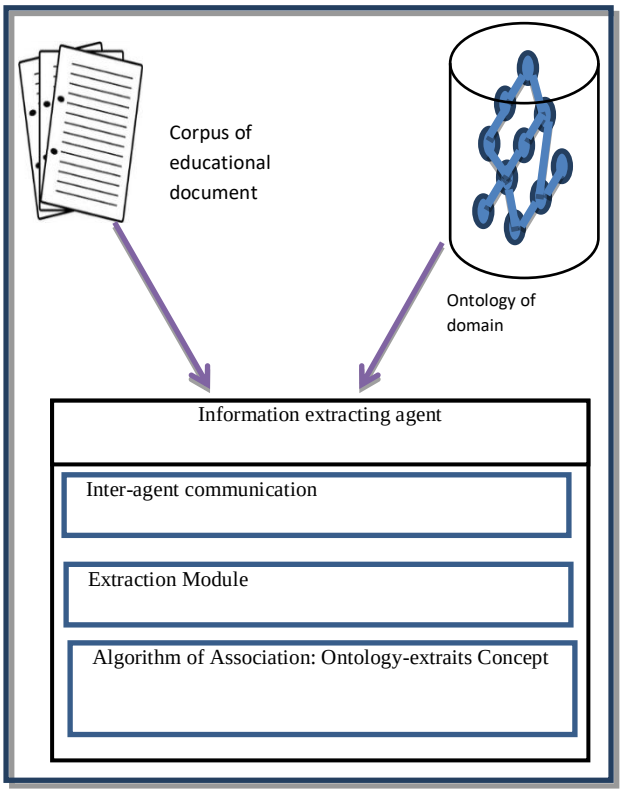

Figure 8: Architecture of extraction agent

With the module of learning, different teaching strategies are represented and control methods and are implemented using an algorithm for sequencing the appropriate strategies. This module supports the promotion of learning through the design, control and organization of training to be conducted for each student. It is also able to determine the concepts that should be available at all times, how to present each concept, and when and how to interrupt the student.

The treatment agent stores information specific to each learner. It follows how a student with the subject matter and provided the relevant answers for each student.

IX. OUR SYSTEM WORKS

Automating the annotation of educational materials can operate a recognition technique of teaching concept to locate and classify these concepts into predefined categories aims. It can also be based on the search terms associated with the concepts of domain ontology.

A. *Operation of the extracting agent*

The objective of this agent is to generate RDF metadata associating the nodes of the document domain concepts described by the terms or PC was located in these nodes. The metadata generated using the parameters defined in the module annotation.

The main steps of this agent are:

- XHTML documents relating to the field are cleaned and passed the extracting
- The approach begins by extracting label documents before applying them a set of

patterns for the terms or PC candidates. We call term candidate (or candidate PC) a term (or an educational concept) excerpt and not yet associated with a concept of the ontology.

- The terms or PC candidates obtained are compared with the concepts of the domain ontology. It is made directly with the terms, the PC or labels of ontology concepts.

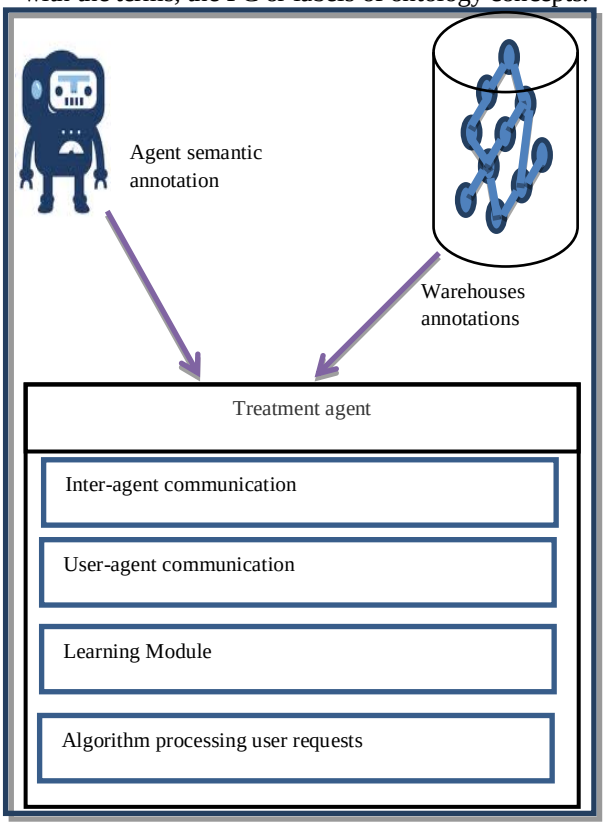

Figure 10: Architecture of the treatment agent

The Information Module Extraction analyzes the content of the document in its XML form, the structure is a DOM (Document Object Model). This action provides a new format in which linguistic labels become XML elements and other available information, such as labels become XML attributes. As for the textual unit labeled, it becomes the value of the XML element. The advantage of this format is that it normalizes somehow all extractions, regardless of the engine which produced and the output format required by the latter.

The module will apply the extraction algorithm to transform the conceptual tree in XML format RDF for semantic annotations. This is necessary to preserve the flexibility and scalability of the system as a whole.

B. Operating of the annotation agents

The purpose of the annotation agent is to annotate nodes documents under annotation module that we have defined. We defined an algorithm to generate all annotations. This algorithm exploits the semantics associated with the nodes by the extraction agent and their structural relationships.

The annotation of a node is defined by the presence of one or more terms or PC partner in this node and concepts with which they were associated. An annotated node is an instance of one or more ontological concepts

Module Annotation is designed to handle different types of content, the text but also video or image, thanks to the coupling of suitable extraction agent. For now, only the textual resources can be treated

A document or a set of documents is submitted to the agent information extraction that enables annotation whose module annotation deals annotate each document agent. The annotation phase is controlled to the concepts defined by the domain ontology. Output, new instances are directly imported into the knowledge base and annotations are stored in files or sent to the warehouse for storage annotation together with the source document. Our application has implemented a semi-automated process; the human user is responsible for validating the proposals at both annotations as instances.

The RDF document which was built at the end of the extraction agent is transmitted to module Semantic Annotation. This module applies the dedicated algorithms for semantic annotation of educational concepts, such as control of the existence of references to instances or descriptors in the database of the application and verification of domain constraints and scope.

This implementation allows modularity and flexibility particularly effective system to meet closer to the needs of an e-learning application.

C. Operation of the treatment agent

The RDF-based annotations and built can be queried using queries. These queries are reformulated with metadata annotation module, from a user request. These reformulations can be constructed by applying different transformation functions. They achieve all the answers sorted by level of aggregation of the terms or PC identified.

To ensure learning that meets the needs of users, we propose a set of resources that the student can access on its own initiative. We choose to use ontologies and semantic annotations to organize training resources. Defined in the learning module, ontological concepts represent concepts discussed and the semantic annotations to structure the set of concepts of the ontology by the links they establish and the semantics attached to each concept. The structure thus created fosters retention of information by placing each concept relative to those related to it and helps the learner to navigate easily within knowledge.

The learning module includes different education as learning design strategies, regulation and organization of training to be conducted for each student, an algorithm is implemented to determine the type of information available at all times, how to present the information, and when and how to assist the student in his career.

X. CONCLUSION

Semantic annotation of educational materials are used to associate elements of document concepts, instances of concepts or instances of relationships which are formally described in ontology. The objective is to formulate queries based on the vocabulary of this ontology to reason about the semantic annotations, and collect responses combining data from different documents.

We are interested in semantic annotation by intelligent agents educational materials for online training related to a domain. This area is described in an ontology that maps each concept in a set of representative equivalent labels and a set of terms or teaching concepts (PC). We

propose a fully automatic approach which aims to define instances of concepts as perfectly as possible and to identify the possible presence of an instance in a document node and annotate this node.

REFERENCES

1. Lenne, D., Abel, M.H., Benayache, A. & Moulin, C. (2005). "E-MEMORAe: an e-Learning Environment Based on an Organizational Memory". In P. Kommers & G. Richards (Eds.), Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2005 (pp. 4651-4658).
2. Neepa K. Shah, "E-Learning and Semantic Web", International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 2, No. 2, April 2012
3. Anand Tamrakar, Kamal K.Mehta, "Analysis of Effectiveness of Web based E-Learning through Information Technology", International Journal of Soft Computing and Engineer ing (IJSCE) Volume-1, Issue-3, July 2011
4. Stojanovic, L., Staab, S. and Studer, R. "eLearning based on the Semantic Web". WebNet2001, World Conference on the WWW and Internet, October 23-27, 2001, Orlando, Florida, USA
5. Ljiljana Stojanovic, "eLearning based on the Semantic Web", FZI Research Center for Information Technologies at the University of Karlsruhe, 76131 Karlsruhe, Germany
6. Berdjouh Chafik, "agent approach for service discovery", Centre de Formation professionnelle El-Meghaier Wilaya El-OUED, ALGERIE, Special Issue on ICIT 2009 Conference – "Web and Agent Systems"
7. Berners-Lee T., "Weaving the Web", Harper Eds, San Francisco, 1998, p.226.
8. Handschuh S. et Staab S. (Eds).. "Annotation for the semantic web". Vol. 96 of Frontiers in Artificial Intelligence and applications", IOS Press. 2003.
9. Sbaa.A, R. El Bejjet and H.Medromi, "Architecture design of a virtualized embedded system ", International Journal on Computer Science and Engineering (IJCSE), Vol. 5 No. 01 Jan 2013.
10. CORCHO O., "Ontology based document annotation: trends and open research problems", in International Journal of Metadata, Semantics and Ontologies, Inderscience, 2006, pp. 47-57.
11. DZBOR M., DOMINGUE J. & MOTTA E., "Magpie – towards a semantic web browser", in Proceedings of the 2nd International Semantic Web Conference (ISWC'03), International Handbooks on Information Systems, Springer-Verlag, 2003, pp. 690-705.
12. DOMINGUE J., DZBOR M. & MOTTA E., "Semantic Layering with Magpie", in Handbook on Ontologies, STAAB S. & STUDER R. (Eds.), Frontiers in Artificial Intelligence and Applications, Volume 96, IOS Press, Springer-Verlag, 2003, pp. 533-554.
13. DILL S., EIRON N., GIBSON D., GRUHL D., GUHA R., JHINGRAN A., KANUNGO T., RAJAGOPALAN S., TOMKINS A., TOMLIN J. A. & ZIEN J. Y., "SemTagand Seeker: Bootstrapping the Semantic Web via Automated Semantic Annotation", in Proceedings of the 12th International World Wide Web Conference (WWW'03), ACM Press, Budapest, Hongrie, 2003, pp. 178-186.
14. DILL S., EIRON N., GIBSON D., GRUHL D., GUHA R., JHINGRAN A., KANUNGO T., MCCURLEY K. S., RAJAGOPALAN S. & TOMKINS A., "A Case for Automated Large-Scale Semantic Annotation", in Journal of Web Semantics, Science, Services and Agents on the World Wide Web, Elsevier, 2003, pp. 115-132.
15. CIRAVEGNA F., DINGLI A., GUTHRIE D. & WILKS Y., "Integrating Information to Bootstrap Information Extraction from Web Sites, in Proceedings of the workshop on Information Integration on the Web" (IJCAI'03), Acapulco, Mexique, 2003.
16. DINGLI A., CIRAVEGNA F. & WILKS Y., "Automatic Semantic Annotation using Unsupervised Information Extraction and Integration", in Proceedings of the Workshop on Knowledge Markup and Semantic Annotation (KCAP'03), Sanibel, Floride, 2003.
17. CIRAVEGNA F., CHAPMAN S., DINGLI A. & WILKS Y., "Learning to Harvest Information for the Semantic Web, in Proceedings of the 1st European Semantic Web" Symposium (ESWS'04), SpringerVerlag, Heraklion, Crête, Grèce, 2004.
18. Popov B., Kiryakov A., Manov D., Kirilov A., Ognyanoff D. & Goranov M. (2003). "Towards semantic web information extraction". In Proceedings of the Human Language Technologies Workshop (ISWC'03), p. 1–22, Sanibel, Floride.
19. Amardeilh F., Laublet P. &Minel J.-L. (2005). « Annotation documentaire et peuplement d'ontologie à partir d'extractions linguistiques ». In Actes des 16èmes journée francophones d'Ingénierie des Connaissances, p. 25–36.
20. AUSSENAC N. & SEGUELA P., ‡ Les relations sémantiques : du linguistique au formel, in Cahiers de grammaire ‡, Numéro spécial ‡ Sémantique et Corpus ‡, Volume 25, Presses de l'UTM, Toulouse, 2000, pp. 175-198.
21. Martins B., Silva M.J., ‡The WebCAT Framework - Automatic Generation of Meta-Data for Web Resources‡, Web Intelligence 2005, Compiegne- France, p. 236-242.
22. Shchekotykhin K. M., Jannach D., Friedrich G., Kozeruk O., ‡ AllRight: Automatic Ontology Instantiation from Tabular Web Documents ‡, The 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference, 2007, Busan-Korea, p. 466-479.

AUTHOR PROFILE

Aziz Oriche is with the computer science department, Faculty of Science Abd El Malek Essaadi, Tetouan, BP: 93002 Morocco

Abderrahman CHEKRY is with the computer science department, Faculty of Science Abd El Malek Essaadi, Tetouan, BP: 93002 Morocco

Mohamed KHALDI Professor NUS (normal upper school) martil, Tetouan, BP: 209 Morocco