

# Web Resources Annotation for the Web of Learning

Jawad Berri

Information Systems Department  
King Saud University  
Riyadh, Saudi Arabia

**Abstract**—Semantic annotation of web resources is an essential ingredient to leverage the web of information to the semantic web where resources are easily shared and reused. In the education field, reusing hypermedia web resources can support to a great deal the design of modern instructional environments and the development of interactive and non-linear material for learning. Sharing and reusing these resources by different web applications and services presupposes that they are visible for retrieval through a semantic description of their content, function and relations with other resources. This paper presents the annotation and discovery of web resources to create learning objects that constitute the building blocks of learning sessions which are delivered to users in the Web of Learning. Semantic annotation is done by the contextual exploration method which analyzes web resources' text descriptions and metadata in order to annotate automatically resources. We present the system architecture and a case study that illustrates the proposed approach.

**Keywords**—Web resources; semantic annotation; web of learning; contextual exploration

## I. INTRODUCTION

Data has proliferated on the web during the last decade resulting in a huge amount of web resources. Web 2.0 technologies eased this information rise by providing tools for collaboration and sharing. The advances in mobile technologies has also facilitated for users to produce and upload with few clicks web resources that are conveniently shared on the web. Sharing and reusing these resources by different web applications and services presupposes that they are described semantically.

This task is a necessity to mutate the existing web of information to the semantic web (or Web 3.0) [1, 2]. Semantic description allows a web resource to be searched and retrieved in accordance with its substance, the function it achieves and its relationships with other resources. It will be visible through its semantic description and not simply keywords, which makes it easy for semantic web search engines to discover it. In the education field, reusing hypermedia web resources can support to a great deal the design of modern instructional environments that exploit available information and ubiquity of technologies. Hypermedia web resources such as video and audio files, images, wikis, presentations, web documents and others are particularly interesting as they allow the development of interactive and non-linear material for learning. Reusing hypermedia resources for learning will also relieve educators from the burden of systematically authoring learning material which is a major bottleneck in the design of instructional environments.

## A. Intuitive Instructional Environments

The development of new instructional environments is a necessity as learners in the age of technology are exposed all the day to different kinds of sophisticated devices and very rich information content. They interact with their devices with ease and have intimate relations with them. In this hi-tech environment characterized by rich content hypermedia information, learners are expecting to be exposed to familiar environments when they seek information, communicate, play games and learn. Although lots of efforts have been deployed to develop learning environments in the field of education by using Web 2.0 technologies, there is still work to be done to create learning spaces where learning becomes intuitive and more adapted to the real needs of learners. In fact, information should be disseminated in such a way that users looking for needed information become learners as they will be able to deepen and diversify their knowledge through durable learning. Making use of available hypermedia web resources can contribute greatly towards the creation of intuitive learning spaces as they promote interaction and allow fluent navigation over the learning environment.

Web 2.0 technologies and hypermedia information available on the web can contribute greatly in the education field. Web resources can be reused and aggregated with other material to fit education purposes. Although the majority of web resources available are not meant to be used for education, consulting and viewing these resources by users is a form of learning as knowledge is acquired and used in their lives. For example a video which shows how wild life animals hunt a prey can be used into a biology class to illustrate the food chain principles. Users accessing the web seeking information is a form of ad hoc learning requiring the user to spend time looking for adequate resources that allow him/her to build relevant and enough knowledge about a topic of interest.

## B. The Web of Learning

The *Web of Learning* is a learning ecosystem build on the top of the existing web (the web of information) which makes use of existing web resources and organizes them to fit education purposes [3]. It aims to reorganize web information in order to provide users with learning spaces that are generated from information and web resources. The Web of Learning is characterized by a set of features: i) the use of hypermedia resources available on the web that match the learner's needs; ii) the integration of different forms of hypermedia information to involve diverse cognitive human activities in learning; iii) it allows learners to construct their personal learning pathways among the proposed learning structure to promote active and adaptive learning; and iv) it manages learning through sessions

allowing learners to learn according to their pace and constraints and can resume their learning at any time. Web of Learning promotes just-in-time-learning which means that the generation of learning material is done when needed. This paradigm supports education on demand and allows users to control the pace and course of learning. In this paper we address mainly the annotation of hypermedia web resources as a requirement to reuse them in the web of learning. We present also the semantic annotation architecture which creates learning spaces on-demand. The architecture relies on the annotation and discovery of web resources through their semantic description to generate learning objects that constitute the building blocks of learning sessions which are delivered to users in the Web of Learning. Semantic annotation is done by the contextual exploration: a linguistic method which analyzes web resources' text descriptions and metadata in order to annotate them automatically.

This paper is organized as follows: next section presents similar approaches that have been proposed in semantic annotation. Section 3 explains how automatic semantic annotation of resources can serve the education field. Section 4 presents the contextual exploration method as a computational system that can carry out resource annotation automatically. Section 5 exposes the system architecture and the following section illustrates the approach through a case study. In the last section we conclude this work and propose research work to be undertaken in the near future.

## II. RELATED WORK

Research about Semantic Annotation is abundant specifically with Web 2.0 wave which have favored social interactions between web users. Social tagging or social annotation is the particular activity by which a web user associates a semantic tag (text string) to a web resource (video, image, web page, web application, web service, etc.). Most of the Web tools and systems offering this facility do not restrict the tag selection and give the user the freedom to associate any tag to any resource. Some others use a controlled vocabulary or a thesaurus from which the user chooses the tag. Semantic annotation is very practical for organizing web resources as it improves the search of resources on the web and provides useful recommendations based on the resource content and correlations between resources, users, interests etc. sharing similar tags. In [4] Andrews et al. present a semantic annotation classification of tools and models based on three criteria: i) the structural complexity which indicates the amount of information associated with the annotation. Four types of annotations are proposed based on their complexity: *tags* describing a particular resource property such as the name of a place in a picture, *attributes* which define a property of the annotated resource such as location or starting date; *relations* which relate a resource with another one; and *ontologies* which allows to describe the resource with respect to an ontology whereby relations, properties, and restrictions will hold among resources.; ii) the vocabulary type which describes the vocabulary used for annotation: free-form natural language text, controlled vocabulary or ontology; and iii) the user collaboration (single-user or community models) which denotes the way users contribute to create different types of annotations. Zubiaga et al. proposed in [5] an approach to

produce an automated classification of resources based on existing social tag sets from the social tagging systems: Delicious<sup>1</sup>, Library Thing<sup>2</sup> and Good Reads<sup>3</sup>. The approach uses the Support Vector Machines classification algorithm to derive two main results: the way users tag, and the way tags are used by the social tagging system to automatically classify resources. The success of the approach is mainly dependent on factors related to the tagging system itself such as whether the tags are suggested by the system or freely added by the user. Yu and al. in [6] present an approach to annotate educational video resources for distance learning. The approach uses the Linked Data [7] technology and ontologies to annotate videos. The authors developed a tool (Sugar-Tube) that searches resources based on the annotations associated to videos. Moreover, the tool allows to link videos with other educational resources from the Linked Open Data cloud and the web. Another work proposed by Lau and Lee [8] considers annotating educational resources using social tags. The authors present a system that uses folksonomy tags to filter, rank and recommend learning resources which are annotated to fit with the users' needs on a social learning environment. Smine et al. [9] propose a semantic annotation approach based on the contextual exploration method to annotate learning objects. The tags are then used to create a semantic inverted index in order to retrieve learning objects.

Much research has been dedicated to web resource annotation for reuse and sharing. In this paper we are interested to annotate web resources for use in a learning environment. Our approach considers the web as a learning space where resources can be annotated to fit with specific learning scenarios for users in context.

## III. SEMANTIC ANNOTATION

Semantic Annotation is the process that associates attributes, comments, descriptions or any other metadata to a resource. This task represents one of the objectives stated by the semantic web initiative of having data on the web defined and linked in such a way that it can be used by machines for automation, integration and reuse across various applications [10]. Lots of works [11] have been done in this direction led by the W3C initiative<sup>4</sup> resulting in the definition of web standards [12] for semantic annotation (RDF, FOAF, etc.) and tools (Semantic Media Wiki [13], Annotea [14], KIM [15]) which allow semantic annotation of resources. These efforts have settled a web environment ready for dealing with semantic annotation of resources where users can post their authored resources along with semantic descriptions for public use; however, less focus has been dedicated to develop computational models, architectures and tools that can annotate automatically web resources.

Automatic annotation of web resources is an important web ingredient to leverage existing information and to foster the web mutation. It is essential for resource discovery and reuse. Resources will be produced and then described automatically

---

<sup>1</sup><http://delicious.com>

<sup>2</sup><http://www.librarything.com>

<sup>3</sup><http://www.goodreads.com>

<sup>4</sup><http://www.w3.org/standards/>

by semantic annotation tools in order to be reused in many fields and by different web services and applications. The development of semantic annotation tools that can annotate automatically web resources has many advantages: i) it decouples resources authoring from their semantic description and use. Resources can be produced for a specific purpose and can be reused by applications and services for other purposes if their semantic description fits with the application needs and context; ii) it facilitates resources reuse and sharing as resources will be visible through their semantic description which makes it easy for semantic web search engines to discover them; and iii) it will relief authors from the burden of describing all the features of their resources in detail. In this work we present a semantic annotation system that is able to annotate resources automatically from their text descriptions.

### C. Web Resources

The concept of *Web Resource* is fundamental in the web architecture. It is the primitive element that constitutes the web and that can be addressable. A web resource is identified by its Uniform Resource Locator (URL) which has been used originally to address documents and files on the web. The concept has evolved with the diversity of schemes for web resources to encompass any “entity” or “thing” that can be identified in a networked information system. Therefore, the identification of web resources has been extended to Uniform Resource Identifier (URI) specification which provides a simple and extensible means for identifying a resource [16]. Web resources can be described semantically using the Resource Description Framework (RDF)<sup>5</sup> language. RDF is based on XML which facilitates processing, exchange and reuse of web resources and their associated descriptions.

### D. Web Resources for Learning

Using web resources in learning is an interesting field to investigate. Modern instructional environments such as those developed in E-Learning or M-Learning are using Hypermedia Web Resources (HWR) embedding multimedia and hypertext medium of information to create interactive non-linear material for learning. The availability of multimedia development software (such as animation tools, presentation tools, web authoring tools, and others) has facilitated the production of HWR such as videos, audios, images, presentations, web documents, maps, news, emails, web services, mobile applications, wikis, blogs, podcasts, etc. HWR are very suitable for learning as they provide more control and adaptation on the learning flow. They also promote interaction and collaborative learning as they ease navigation over the instructional environment. In this research we focus on the description of HWR for resources annotation and discovery to create integrated learning objects. These learning objects are the building blocks of learning sessions that are delivered to users in the Web of Learning.

### E. Web Resources Annotation

Automatic annotation of HWR is not a simple task considering the masses of resources available on the web and the difficulty of apprehending these resources for semantic analysis. This task is challenging the research community

requiring efforts from many disciplines to describe semantically web information. On the technical side, the network infrastructure needs an extension in terms of storage and access to be able to store semantic descriptions and retrieve them with no delay. The design and implementation of effective semantic annotation systems is fashioned by the resource content to analyze for annotation. The digital content of HWR is naturally the most appropriate data to analyze as it contains the essence of the resource. For instance, the analysis of video files involves video content analysis and an interpretation of the scene data [17]. It requires the extraction of contours and features, contrasting colors between regions and comparing frames in order to identify objects, individuals and motion [18]. This field of research is focused on analyzing the physical data of the resource which are the image pixels represented as bits. Tremendous research efforts have been deployed in this area resulting in the development of sophisticated algorithms and software tools that are able to detect for instance human intruders in video surveillance systems, recognize car plate numbers in traffic monitoring systems, kinetic based video games, etc. Although these systems and tools can recognize objects and their motion with precision they cannot infer basic semantic features from a scene in real life such as names of individuals, their roles, the relation between objects, etc. [19]. These semantic features are essential ingredients for semantic annotation. Relying only on the resource’s digital content analysis does not help much in semantic annotation.

Web resources include generally other data added by the resource creator or by users who viewed the resource. This information can be very rich including the name of the resource, a text description, hyperlinks to other related resources and tags associated to the resource. Also when the resource triggers a specific interest, users add their own comments that can be very interesting to analyze for semantic annotation. In addition, social interaction generates metadata that is added to describe the relevance of the resource such as the number of views, the percentage of users who liked the resource, and personal tags that might be added by users. In this work we focus on analyzing the text and metadata related to the resource for semantic annotation. Analyzing text data requires linguistic analysis tools which need to be robust and flexible in order to deal with unrestricted text on the web. Moreover, dealing with resource metadata necessitates decision models to sort resources based on relevance. In the next section we present a linguistic method – the Contextual Exploration Method, that is able to annotate automatically resources based on the text associated with resources and we present a system that has been developed to annotate videos on the web.

## IV. CONTEXTUAL EXPLORATION

The Contextual Exploration Method (CEM) can significantly contribute to web resource annotation. CEM is a computational linguistics method that is suitable to analyze unrestricted text. CEM is the result of many years of research which has led to the development of a framework which has been applied to solve many problems related to language processing [20, 21, 22]. It is a decision-based method that involves grammatical and lexical knowledge regarding a decision making task when solving a linguistic problem. The

<sup>5</sup><http://www.w3.org/TR/rdf-schema/>

method simulates the behavior of a human who is reading a text to analyze it in view of taking a decision. CEM scans the linguistic context looking for linguistic markers that can trigger decisions. Markers can be any word occurrence, morpheme, lexeme, or lexical unit. Once a marker is found, then the context is further analyzed to find linguistic contextual clues surrounding the marker to support taking an unambiguous decision. Linguistic markers and clues are organized into a database and are used by decision rules which annotate text passages. CEM is a flexible and robust method that can deal with web resources repositories which include resources that are freely described by users. Unlike classical natural language processing architectures, CEM does not rely on rigorous parsing and language dictionaries; instead, it uses the linguistic expertise related to the problem at hand and is able to take decisions when annotating texts by mean of heuristics and strategies. CEM can deal with the inherent variations of texts associated with web resources. These texts are written by users who use an open language that is influenced by user's social cultural factors such as the community to which the user belongs, the age, the region and specific customs. For instance, Arabic speaking mobile users refer to the BlackBerry Smartphone by its initials (البي بي *The BB*). This particular naming is used in general by young Arabic speakers. Arabic speaking fans of the Barcelona Football Club refer to their club using many community specific words such as (البارشا، بارسا، البرشا *Albaarshaa, Baarsaa, Albarshaa*) [23]. Social factors play a major role in the expression of the resource description and any social interaction related to the resource; they must be taken into account to process data associated to web resources for semantic annotation.

The Contextual Exploration Method involves a set of resources and development tools to develop contextual exploration modules and knowledge components to annotate web resources. Figure 1 shows the different components involved in developing a Contextual Exploration Module for semantic annotation.

#### F. Resources and Development Tools

The Linguistic Expertise is represented in the knowledge base; it includes all the linguistic markers and clues which are first detected in the text to be analyzed. Linguistic markers and clues are organized into equivalence lists that are invoked by decision rules. Decision rules check the presence of a specific marker and clues inside a particular context consisting of a text passage and accordingly assign a tag. Decision rules are hierarchically organized in the knowledge base in order to solve inherent languages' ambiguities that are due to polysemy or equivocal contexts.

Linguistic tools are all the tools that are essential for preparing texts to be analyzed. This component includes preprocessing, tokenizing, stemming and morphological tagging tools to handle texts and to prepare them for annotation by the contextual exploration module. Preprocessing the text is a necessary initial step which objective is to remove noise and filter out data that is not used in the linguistic analysis such as special characters, Xml tags, encodings, etc. Tokenizing splits the text into tokens and associates to each one a set of data such as its offset position in the text, its sentence, and other data

related to the text physical structure. Stemming allows to reduce words into a canonical form so that to avoid the differences due to word affixes. Stemming is a light process that does not require a language dictionary as words are simply chopped resulting in word stems. This rough process is suitable sometimes when there is no need for words' part of speech and morphological categories. In case the latter information is needed, then contextual exploration uses Morphological Tagging tools to associate morphological information to the text tokens.

The Corpus is the linguistic repository that helps acquiring and validating the linguistic expertise; it is mainly used in the knowledge acquisition phase to gather linguistic markers and clues and to identify the decision rules.

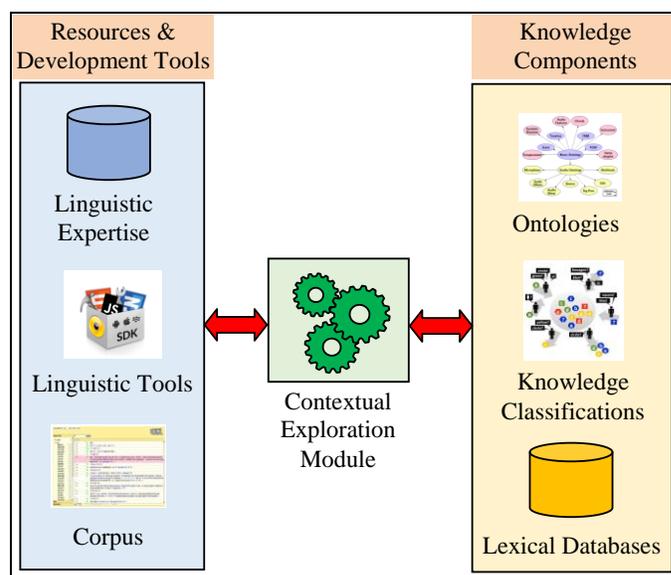


Fig. 1. Contextual Exploration for Semantic Annotation.

#### G. Knowledge Components

CEM requires a specification of knowledge related to semantic annotation. This is done through knowledge components namely: ontologies, knowledge classifications and lexical databases. Ontologies are formal descriptions of concepts, their properties, types and relationships holding between them in a domain of knowledge. When semantic annotation is about knowledge domains for which an ontology is available, CEM uses this ontology in order to annotate texts. For instance in the linguistic field, CEM used the tense and aspect ontology as defined by [20]. When an ontology is not available, which is the case when we deal with open web repositories and unrestricted text, CEM uses available knowledge classifications that are set by domain experts such as taxonomies, thesauri, controlled vocabularies, or folksonomies which generally emerge from web social interaction. Moreover, Lexical Databases are useful to enrich existing knowledge classifications. Word Net [24] is an example of a lexical database that includes many semantic relationships such as synonymy, hyponymy, hypernymy, and antonymy to extend existing knowledge classifications and to relate existing concepts.

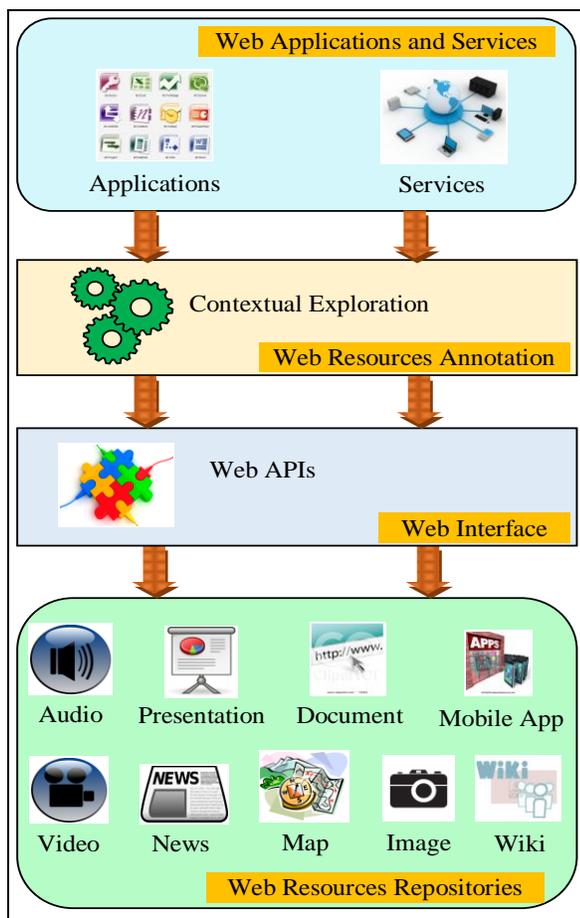


Fig. 2. System architecture

### V. SYSTEM ARCHITECTURE

Figure 2 presents the system architecture for web resources annotation. The architecture includes four components namely: Web Applications and Services, Web Resources Annotation, Web Interface and Web Resources Repositories. Web resources annotation is a process that is triggered by a request from a web application or a web service which seek to use web resources for a specific purpose. The request is analyzed by the Web Resources Annotation component which activates a contextual exploration module specialized in annotating web resources as requested by the web application or service. CEMs include linguistic expertise for annotating specific semantic categories. The activation of the right CEM is done through a fine analysis of the request [23] which detects what type of annotation is required and the objective of the query and then selects the suitable CEM which analyzes the resource’s data and annotates the resource. Accessing web resources requires specific Application Program Interfaces (APIs) to search and get relevant resource’s data for annotation. Most of the popular web hypermedia repositories offer convenient web APIs publicly available to search and retrieve hypermedia resources. The Web Resources Repositories Component represents the available hypermedia resources’ repositories available on the Web. These repositories include a variety of web resources such as video and audio files, web documents, images, presentations, news, wikis, maps, mobile applications, web

services, podcasts, blogs, etc. A first prototype implementing some principles of the Web of Learning has been developed on a mobile platform [25]. The prototype generates learning objects for users on a mobile platform.

### VI. CASE STUDY

In order to illustrate how CEM annotates web resources let’s consider the following scenario: Ayoub is a college student who is interested to learn how to graph quadratic functions. Although he has studied at school a full chapter on this topic, he would like to have a short and concise hypermedia presentation on his tablet computer about this topic. Ayoub submits the following query “How to graph quadratic functions” to the system which is first analyzed and forwarded to many hypermedia web repositories in order to search and retrieve relevant hypermedia web resources corresponding to the information requested. Contextual exploration analyzes the user’s query in order to annotate it and extracts relevant words and phrases. The query objective is detected from the words “How to” hence the query is annotated as (*Objective = method*) which means that the user is looking for a method, a manner or a description of how to accomplish the specific task related to “graph quadratic functions”. Accordingly, information to be retrieved should have the same objective tag so that to match the query. Table I shows part of the linguistic expertise that is used to annotate the query as “method”. In the first column some linguistic expressions denoting the semantic category “method” are represented in the EBNF (Extended Backus-Naur Form) notation [26] where lower case words are terminals and capitalized words are non-terminals. For instance the non-terminal *Lmethod* refers to the list of words expressing a method such as: *method, technique, way, manner, practice, approach and procedure*. The non-terminals “Vcan” and “Vbe” represent all the possible morphological variations of the verbs “can” and “be”. Other semantic categories are used for objective annotation such as: “definition” (“What are mobile agents?”), “cause” (“Why the screen is dark?”), “time” (“When Mona Lisa has been paint”) and “location” (“Where can I store my files?”).

Matching between the query and the resource is done through similar annotations. CEM analyzes the resources’ titles, descriptions and metadata. Resources that are annotated similarly are retrieved and used in the learning object (LO). For some resources the comments left by users may also be analyzed. The linguistic expertise presented in Table I is specific for analyzing queries. It is different from the linguistic expertise to annotate texts present in the descriptions of resources.

TABLE I. LINGUISTIC EXPERTISE FOR THE QUERY OBJECTIVE

Query Objective		
Linguistic Expression	Example	Annotation
[“how” + “to”]	How to save a file?	method
[“how”+Vcan+Ppronoun]	How can I copy an image?	
[Ppronoun]	I, we, someone,	
[“what”+“Vbe”+Lmethod]	What is the procedure to apply for a research grant?	
[Lmethod]	method, technique, way, manner, practice, approach, procedure.	

The following resources have been tagged by CEM as fitting with the objective of the query for the task “graph quadratic functions”:

- From Wikipedia<sup>6</sup>, the system extract the table of contents and the first paragraph’s sentence which introduces the topic: “quadratic functions”;
- From Youtube<sup>7</sup>, the first retrieved video is selected as it has been annotated as “method” due to the presence of the following sentence in the video description: “*I outline a little recipe of things to examine when graphing a quadratic function by hand.*”The underlined words have been spotted by CEM as relevant marker and clues for annotation;
- From Yahoo Images<sup>8</sup>, the system retrieves the third image proposed has been annotated as “method”. The following sentence describing the image in the webpage (“online math learning<sup>9</sup>) has been analyzed as denoting a “method”: “*In this lesson, we shall learn how to graph of quadratic functions by plotting points*”. The underlined words have been spotted by CEM.

Resources are searched into the three popular websites: Wikipedia, Google Images and You tube. We have restricted the system to these websites in a first phase but we can extend itby considering additional APIs to access other websites looking for similar web resources. Once resources are annotated and retrieved, the system packages them into a LOas shown in Figure 3.

The learning object LO0 represents the first LO to display. LO0 includes the title on the top that corresponds to the phrase “Graph Quadratic Functions” which has been extracted from the query. LO0 includes a short text description about the topic that has been extracted from Wikipedia webpage. Beside the text the learning object displays an image from Yahoo images corresponding to the topic. In the middle the learning object includes a video that has been retrieved from You tube about graphing quadratic functions.

Moving from a LO to another one is possible from the navigation map represented into the lower rectangles “Main Topics” and Sub Topics” which include hyperlinks to all possible objects in the learning web of Figure 4. Navigation is not constrained on a specific pathway; the user is free to follow the normal sequence of LOs as suggested by the learning web in which case he should use the “Next” button (represented as an arrow).

He can also navigate randomly by clicking on any topic he wishes to view in which case he will be directed to the specific topic. When the user clicks on any topic or the “Next/Previous” button, he requests to view a LO corresponding to the topic in

question. In such case the system generates the LO for the topic/subtopic requested by the user, displays it on the screen and updates the navigation status in the “Main Topics” and “Sub Topics” rectangles.“Main Topics” contain the topics which correspond to the main sections of Wikipedia Webpage. “Sub Topics” displays any sub sections of the currently displayed topic.

For instance, when Ayoub clicks on the topic “4. Graph”, LO4 is generated the same way as LO0 has been generated and is displayed on the screen. In this case “Sub Topics” will include the sub topic “4.1. Vertex” (Fig. 4).

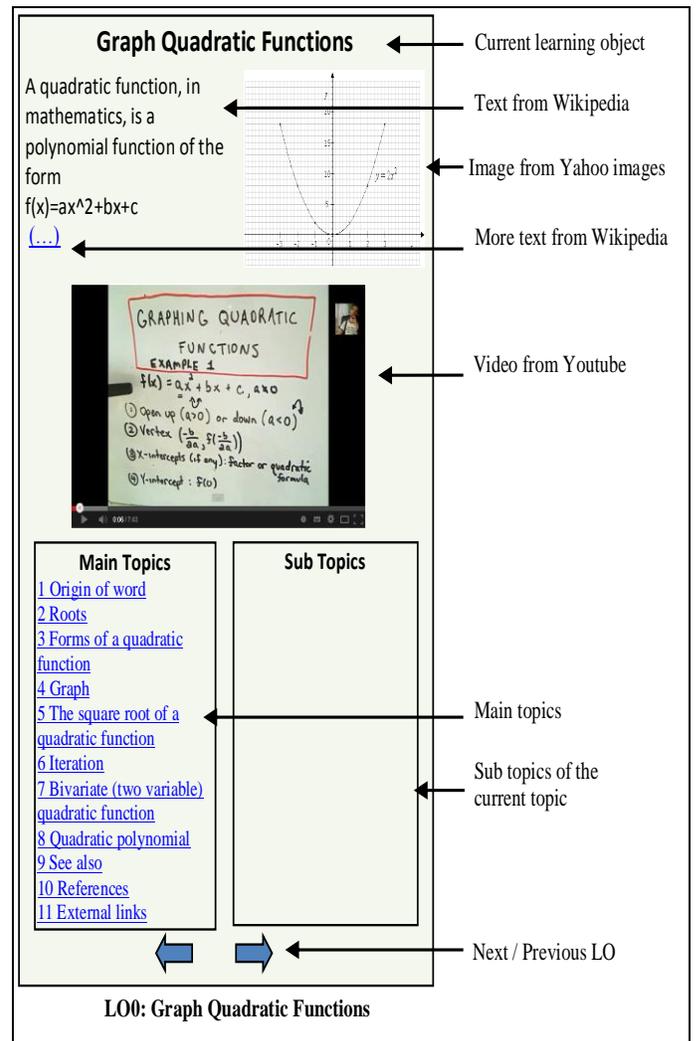


Fig. 3. Learning Object LO0: Graph Quadratic Functions

<sup>6</sup>[http://en.wikipedia.org/wiki/Quadratic\\_function](http://en.wikipedia.org/wiki/Quadratic_function) (accessed on 1st Nov. 2013).

<sup>7</sup><http://www.youtube.com/watch?v=mDwN1SqNMRU>(accessed on 1st Nov. 2013)

<sup>8</sup><http://www.google.com> (accessed on 1st Nov. 2013)

<sup>9</sup><http://www.onlinemathlearning.com/quadratic-graphing.html> (accessed on 1st Nov. 2013)

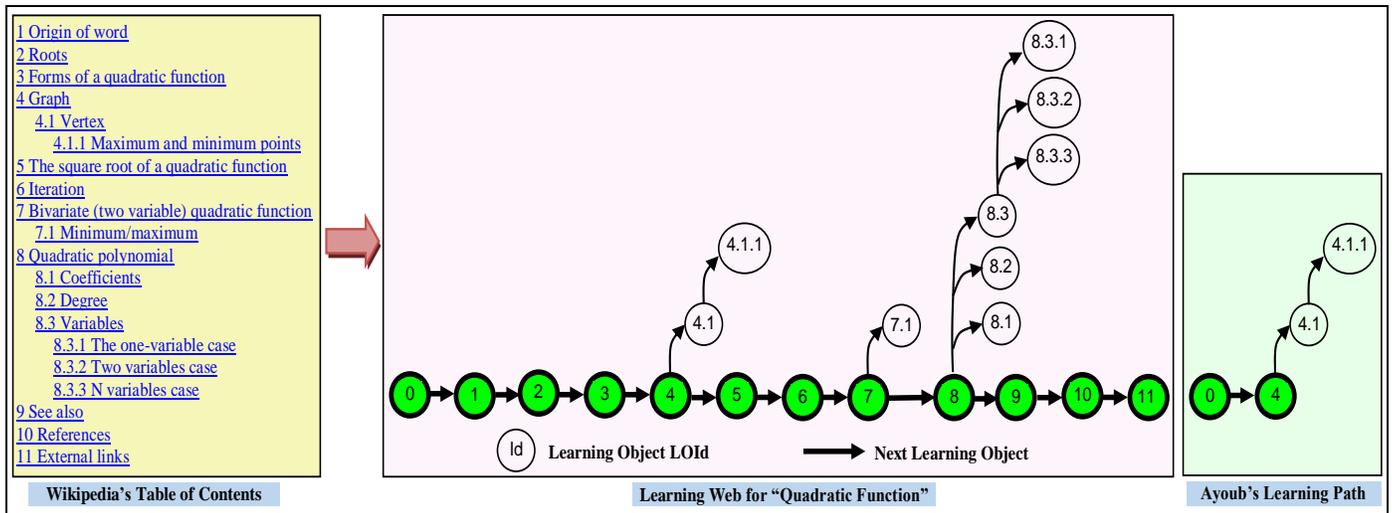


Fig. 4. Learning Web Generation and Learning Path

The system builds the learning web from Wikipedia<sup>10</sup> (Fig. 4). The table of contents of Wikipedia is a good learning structure that organizes the depth and breadth knowledge of a given topic. The Learning Web represents all the possible pathways where the user may navigate in while learning. In our example, Ayoub has visited LO0 then he requested LO4 and as he is interested in Graphing Quadratic Functions. Then he wanted to investigate deeply the topic “4. Graph”, he requested the subtopic “4.1 Vertex” then its subtopic “4.1.1 Maximum and minimum points”. Those LOs requested and visited by Ayoub constitute his personal learning path (Fig. 4 – right graph).

## VII. CONCLUSION

In this paper we presented an approach for reusing available hypermedia web resources to design modern instructional environments that contribute towards the establishment of the Web of Learning. The approach sustains the generation of on-demand interactive and non-linear learning spaces for learners. Prior to be reused, web resources are annotated semantically by the contextual exploration method which analyzes the resources' text descriptions and associates semantic annotations that denote the role and function of the web resource. The implemented system generates on-demand learning material by packaging semantically correlated web resources into learning objects which are plugged into a course map that represents all the possible pathways a learner may navigate in. The system tests done on a set of different topics are encouraging, they show that automatic semantic annotation is more accurate than classical information retrieval in retrieving correlated hypermedia resources [23]. Besides, organizing the learning content into a learning web and allowing users to learn through sessions are much appreciated as they allow users to deepen and diversify their knowledge through durable learning.

Future research aims to extend the linguistic expertise to encompass more semantic categories to have an accurate matching between the user query and the web resources

multimedia descriptions. As for the future extensions of the system, it is important to have a large web coverage by considering more hypermedia websites from where web resources will be considered for annotation and reuse. Furthermore, we would like to develop diversified and more intuitive LO layouts to adapt the packaged hypermedia content to fit the user profile and needs.

## REFERENCES

- [1] T. Berners-Lee, *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web*, HarperCollins, New York, 2000.
- [2] E. Turban, D. King, *Electronic Commerce: A Managerial Perspective*, Prentice Hall, 7<sup>th</sup> Edition, 2012.
- [3] J. Berri, *The Web of Learning: Evolving the Web into a Learning Ecosystem*, submitted for publication, 2013.
- [4] P. Andrews, I. Zaihrayeu, and J. Pane, *A Classification of Semantic Annotation Systems*, *Semantic Web*, vol. 33, pp. 223-248, 2012.
- [5] A. Zubiaga, V. Fresno, R. Martinez, and A. Pérez Garcia-Plaza, *Harnessing Folksonomies to Produce a Social Classification of Resources*, *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, No. 8, pp. 1801- 1813, August 2013.
- [6] H. Q. Yu, C. Pedrinaci, S. Dietze, and J. Domingue, *Using Linked Data to Annotate and Search Educational Video Resources for Supporting Distance Learning*, *IEEE Transactions on Learning Technologies*, vol. 5, no. 2, pp. 130-142, April-June 2012.
- [7] C. Bizer, T. Heath, and T. Berners-Lee, “*Linked Data—The Story So Far*,” *International Journal on Semantic Web and Information Systems*, vol. 5, no.3, pp. 1-22, 2009.
- [8] S. Bounq-Yew Lau, and C.-S. Lee, *Enhancing Collaborative Filtering of Learning Resources with Semantically-Enhanced Social Tags*, 12th IEEE International Conference on Advanced Learning Technologies, pp. 281-285, Rome, Italy, 4 - 6 July, 2012.
- [9] B. Smine, R. Faiz and J.-P. Desclés, *Extracting Relevant Learning Objects Using a Semantic Annotation Method*, International Conference on Education and e-Learning Innovations (ICEELI'2012), Sousse, Tunisia, 1-3 July, 2012.
- [10] J. Hendler, T. Berners-Lee, and E. Miller, *Integrating Applications on the Semantic Web*, *Journal of the Institute of Electrical Engineers of Japan*, vol 122, no. 10, pp. 676-680, October 2002.
- [11] M. Nagarajan, *Semantic Annotations in Web Services*, *Semantic Web Services, Processes and Applications*, J. Cardoso and A. P. Sheth Eds., Springer, pp. 35-61, 2006.

<sup>10</sup><http://www.wikipedia.org>

- [12] L. Causton, Identifying and describing Web resources, European Commission DGXIII/E-4, Interactive Electronic Publishing, 1998. Available at <http://www.elpub.org/brochures/>
- [13] M. Krötzsch, D. Vrandečić, M. Völkel, H. Haller, and R. Studer. Semantic Wikipedia. *Journal of WebSemantics*, 5(4):251–261, 2007.
- [14] J. Kahan, M.-R. Koivunen, E. Prud'hommeaux, and R. R. Swick. Annotea: an open RDF infrastructure for shared web annotations. *Computer Networks*, vol. 39, no. 5, pp. 589–608, August 2002.
- [15] B. Popov, A. Kiryakov, D. Ognyanoff, D. Manov, and A. Kirilov. KIM - a semantic platform for information extraction and retrieval. *Journal of Natural Language Engineering*, vol. 10, no. 3-4, pp. 375–392, September 2004.
- [16] T. Berners-Lee, R. Fielding, and L. Masinter, RFC 3986, Uniform Resource Identifier (URI): Generic Syntax, 2005. Available at <http://tools.ietf.org/html/rfc3986>
- [17] J. K. Aggarwal and Q. Cai, Human Motion Analysis: A Review, *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428–440, March 1999.
- [18] G. Papari, N. Petkov, Edge and line oriented contour detection: State of the art, *Image and Vision Computing*, vol. 29, pp. 79–103, 2011.
- [19] A.K. Jain, L. East, R.P.W. Duin, and M. Jianchang, Statistical pattern recognition: a review, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4–37, January 2002.
- [20] Desclés, J. -P.: *Langages applicatifs, Langues naturelles et Cognition*, Hermès, Paris, 1990.
- [21] J. Berri, D. Maire-Reppert, H.-G. Oh-Jeong, Traitement informatique de la catégorie aspecto-temporelle", *T.A. Informations*, vol. 32, no. 1, pp. 77-90, 1992.
- [22] J. Berri, M. Al-Khamis, Information Exploration Using Mobile Agents, *WSEAS Transactions on Computers*, vol. 3, no. 3, pp. 706-712, 2004.
- [23] A. AlAwajy, J. Berri, Combining Semantic Techniques to Enhance Arabic Web Content Retrieval, *International Conference on Innovations in Information Technology (IIT'13)*, Al Ain, UAE, pp. 141-147, March 17-19, 2013.
- [24] G. Miller, WordNet: A Lexical Database for English. In *Communications of the ACM.*, vol. 38, no. 11, pp.39-41, 1995.
- [25] M. Alzaabi, J. Berri, and M. J. Zemerly, Web-based Architecture for Mobile Learning, *International Journal of Informatics*, vol. 3, no. 1, pp. 213–222, 2010.
- [26] R. W. Sebesta, *Concepts of Programming Languages*, 10th Edition, Addison-Wesley, 2012.